

The Human Visual System



Gordon Wetzstein
Stanford University

EE 267 Virtual Reality

Lecture 5

stanford.edu/class/ee267/



nautilus eye, wikipedia

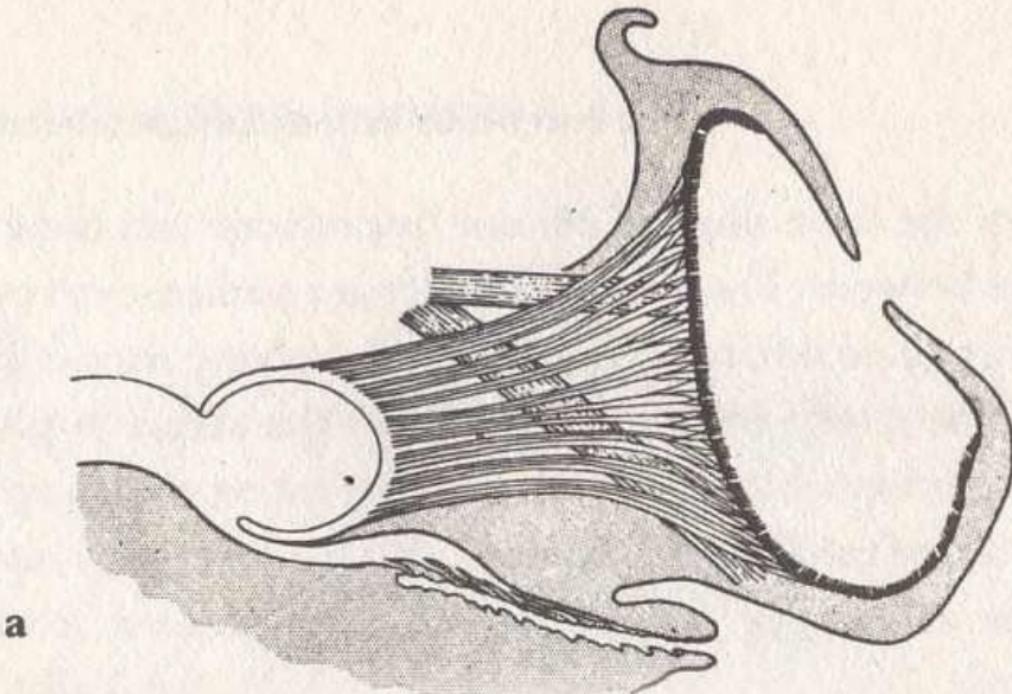
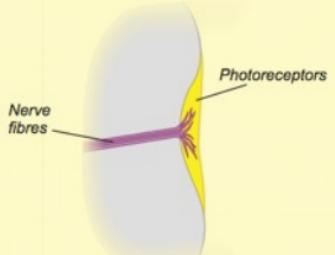


Figure 5.8 (opposite) A range of invertebrate eyes that illustrate approaches to the formation of crude but effective images: (a) *Nautilus*'s pinhole eye; (b) marine snail; (c) bivalve mollusc; (d) abalone; (e) ragworm.

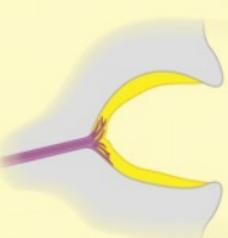


reptile eye, <http://pichost.me/16085580/>

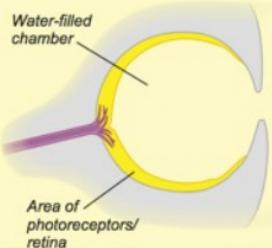
a) Region of photosensitive cells



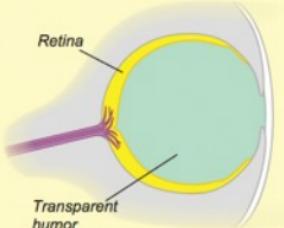
b) Depressed/folded area allows limited directional sensitivity



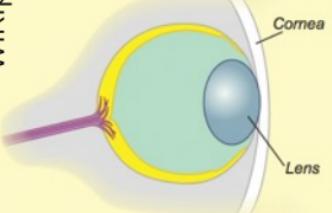
c) "Pinhole" eye allows finer directional sensitivity and limited imaging



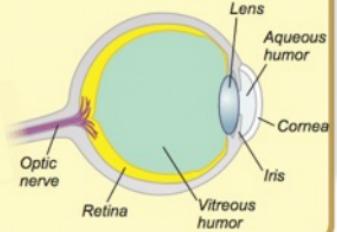
d) Transparent humor develops in enclosed chamber



e) Distinct lens develops



f) Iris and separate cornea develop



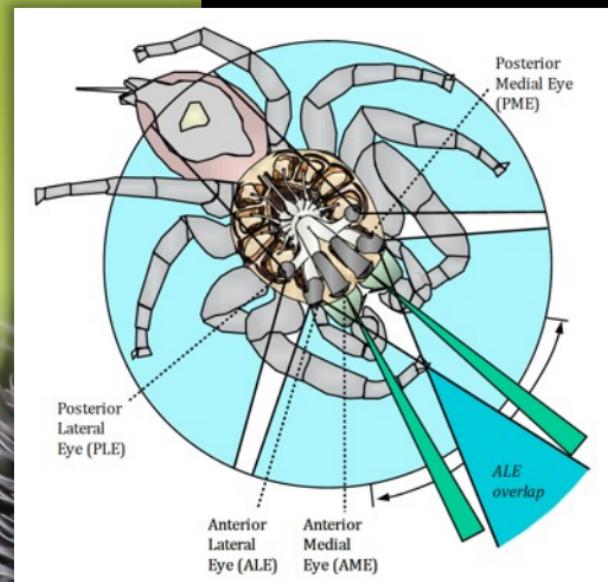
Evolution of the Eye



owl, <https://www.pinterest.com/pin/452400725039917330/>



pigeon, <http://globe-views.com/dreams/pigeon.html>



jumping spider, wikipedia

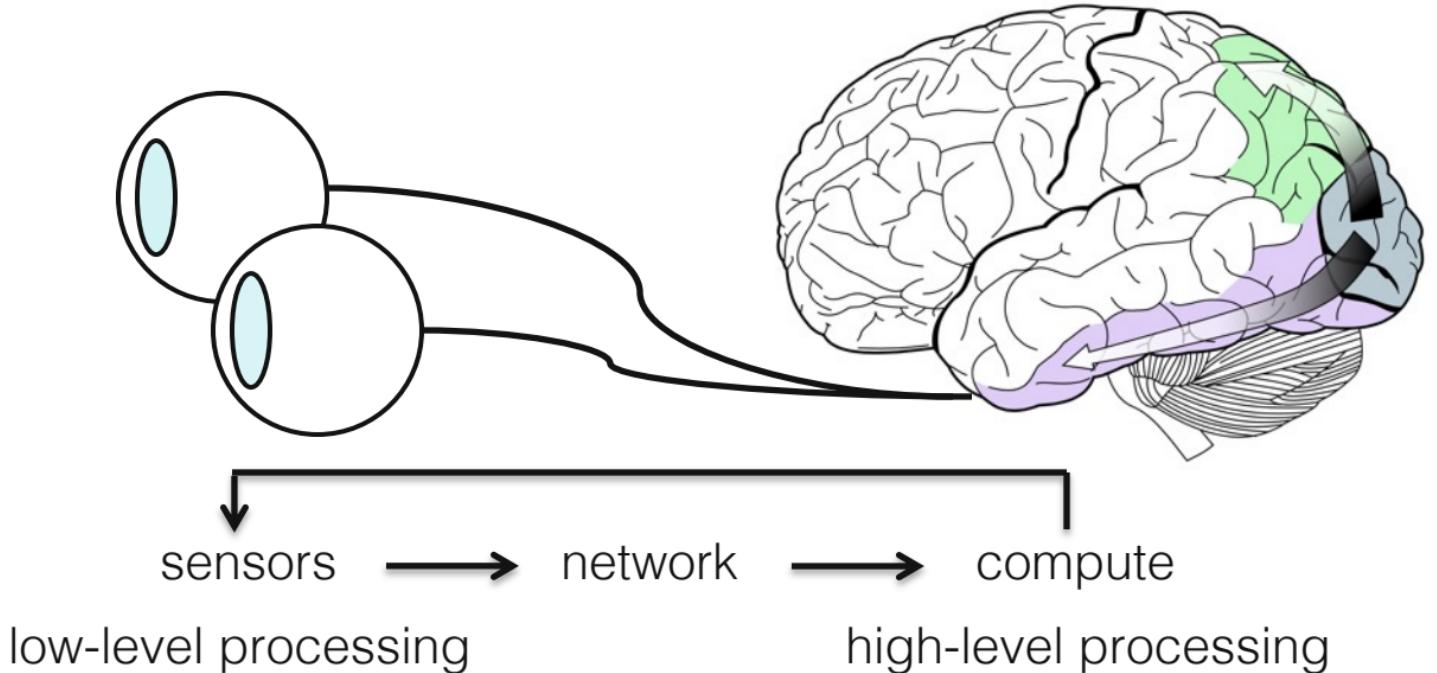


national geographics

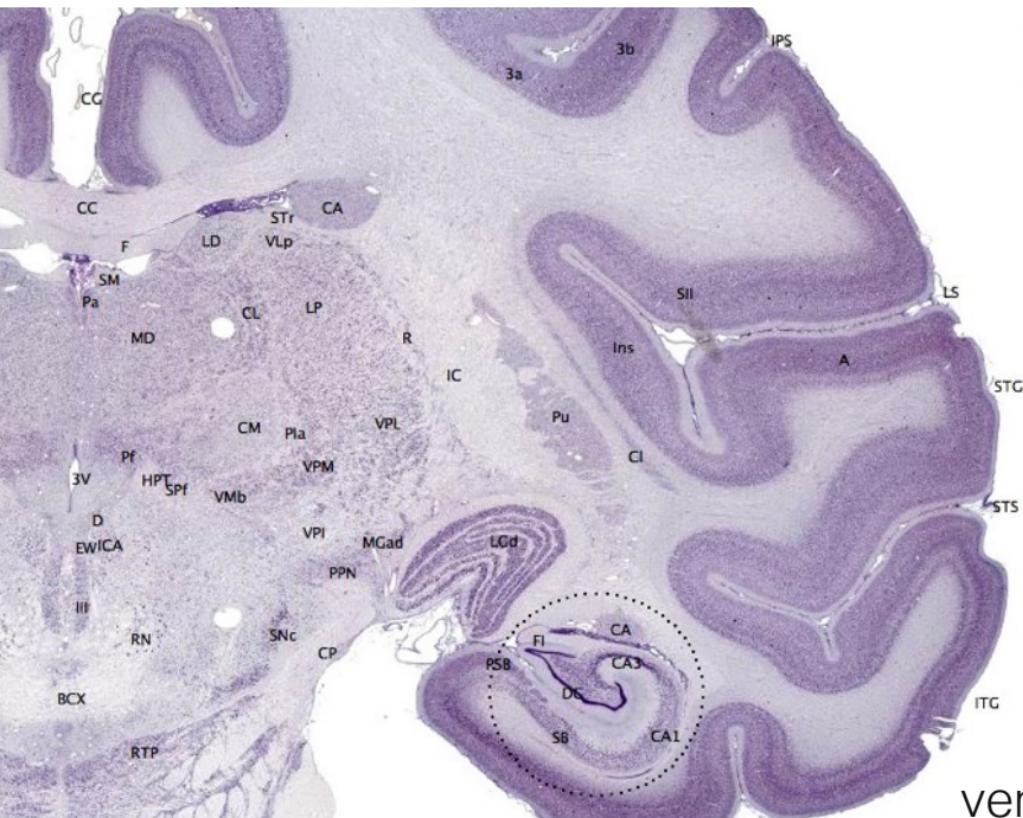
Lecture Overview

- **visual acuity:** 20/20 is ~ 1 arc min
- visual acuity varies over retina: can exploit via **foveated rendering**
- **visual field:** $\sim 200^\circ$ monocular, $\sim 120^\circ$ binocular, $\sim 135^\circ$ vertical
- **temporal resolution:** ~ 60 Hz (depends on contrast, luminance)
- **depth cues in 3D displays:** disparity, vergence, accommodation, blur, ...
- **accommodation range:** $\sim 8\text{cm}$ to ∞ , degrades with age

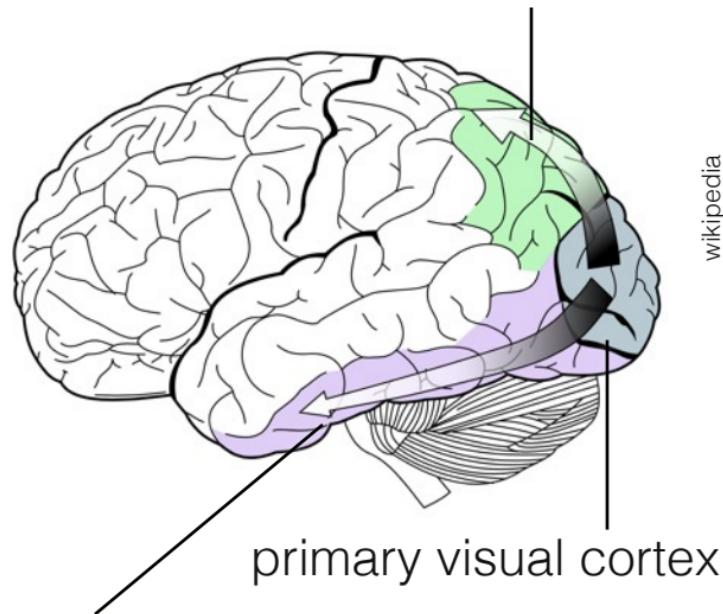
Overview



Overview



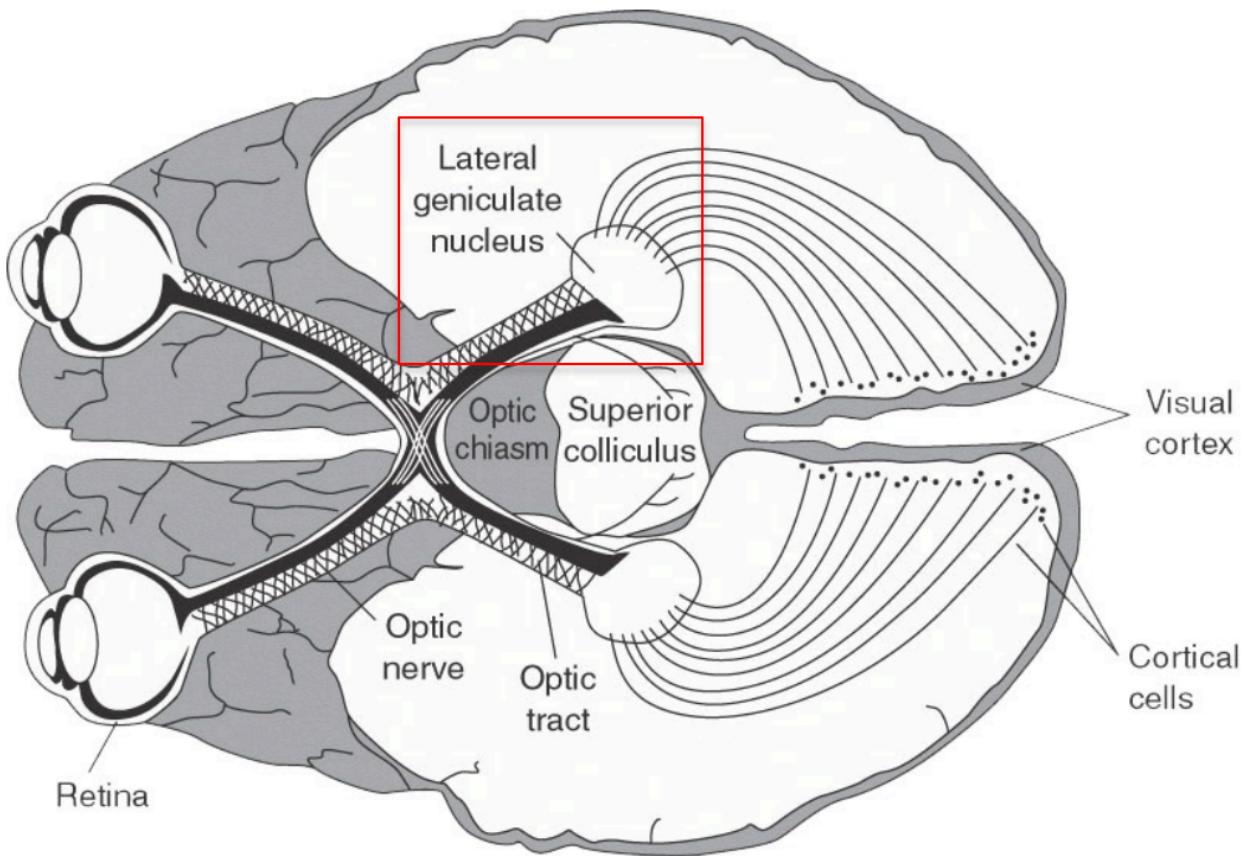
dorsal stream: spatial awareness



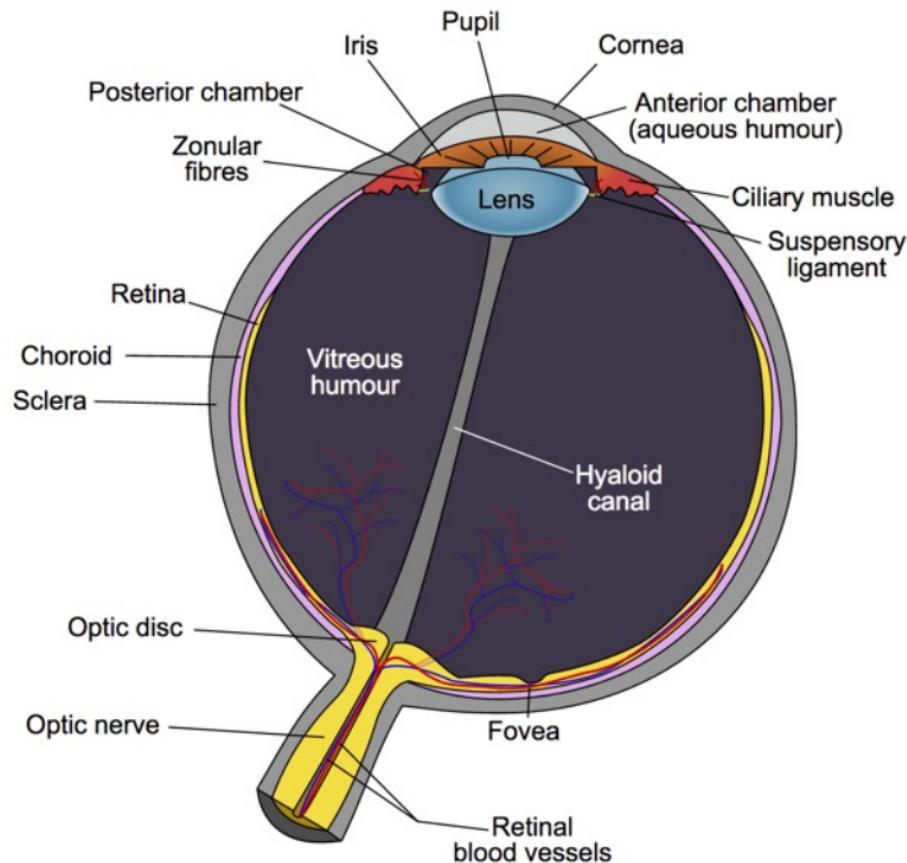
wikipedia
primary visual cortex

ventral stream:
recognition, object identification

Overview

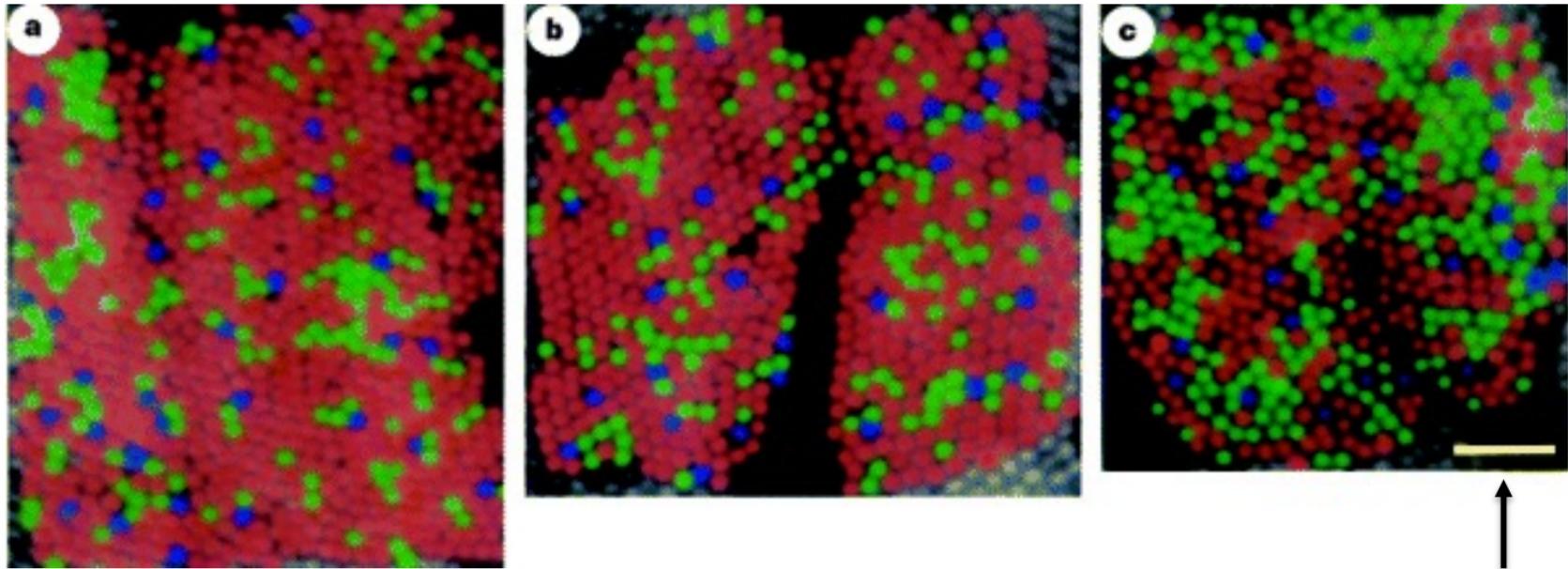


Anatomy of the Human Eye



The Retina

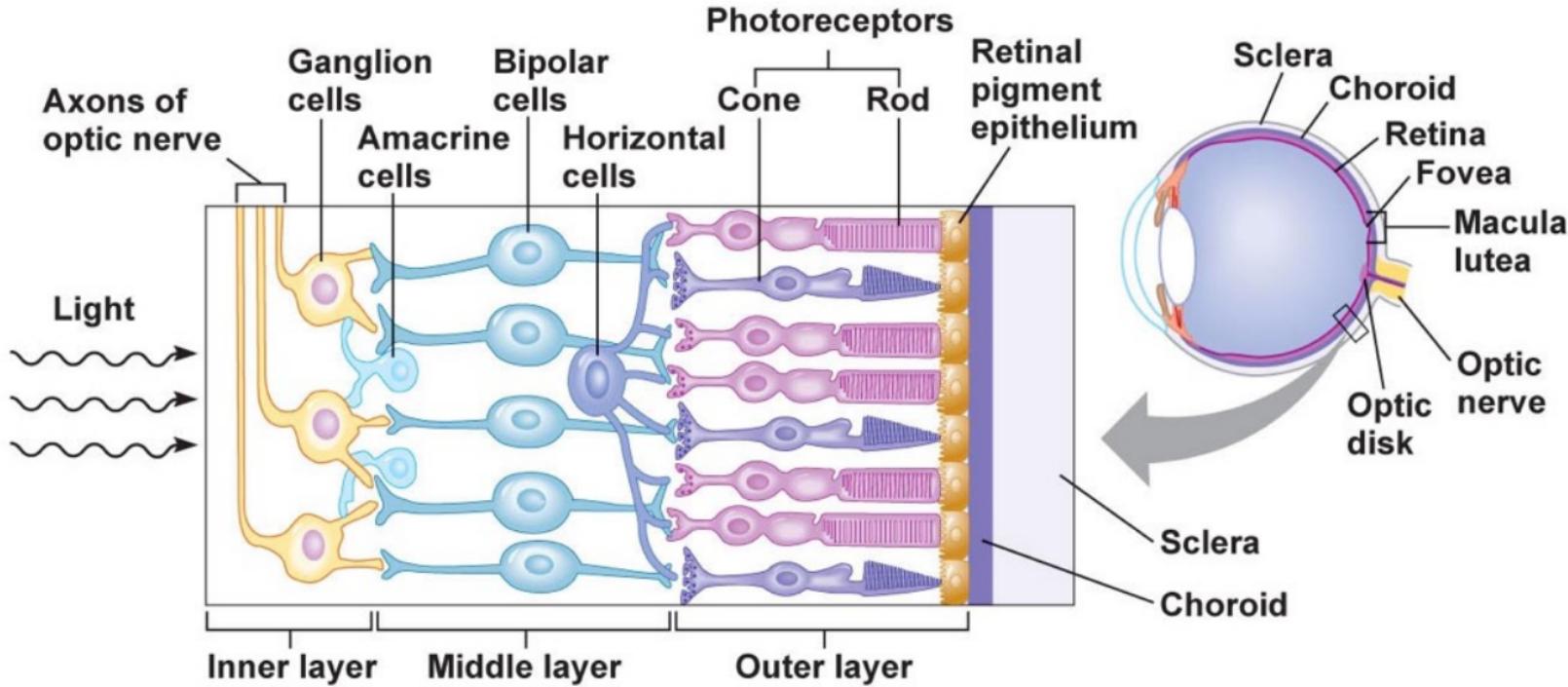
Roorda & Williams, 1999, Nature



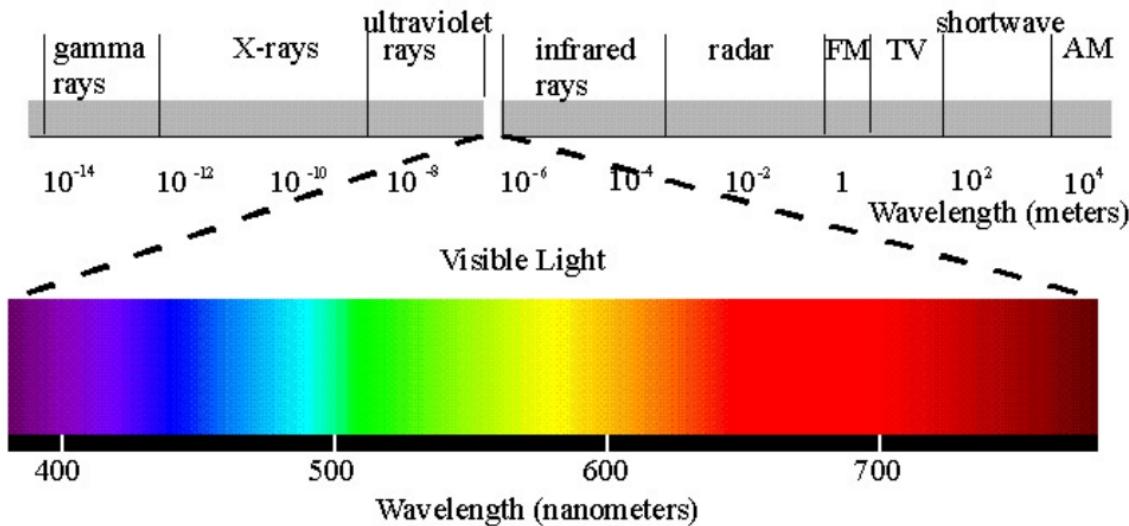
5 arcmin visual angle

photoreceptors: 3 types of cones (color vision), rods (luminance only, night vision)

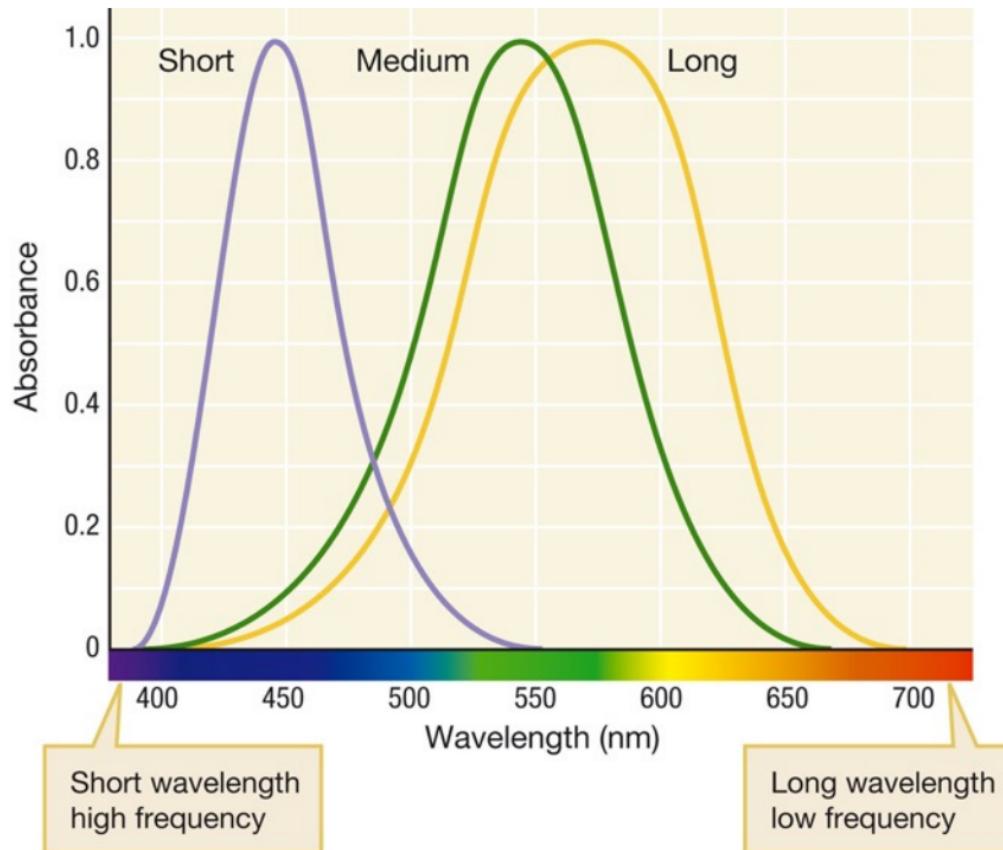
The Retina



Color Perception

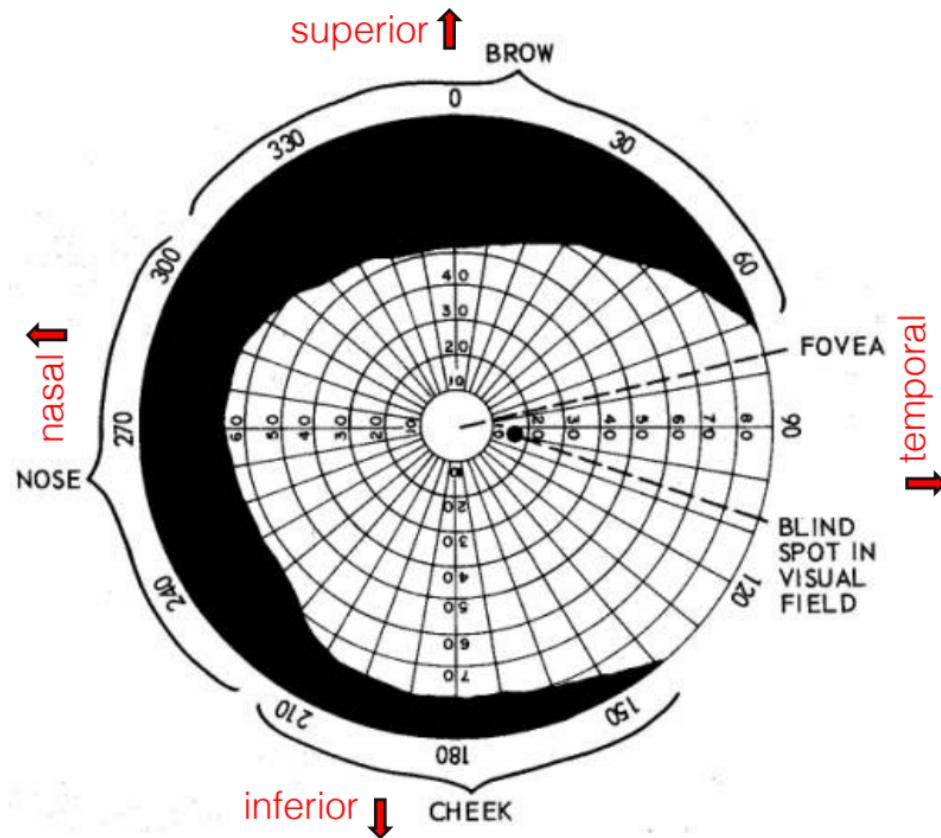


Color Perception - Sensitivity of Cones

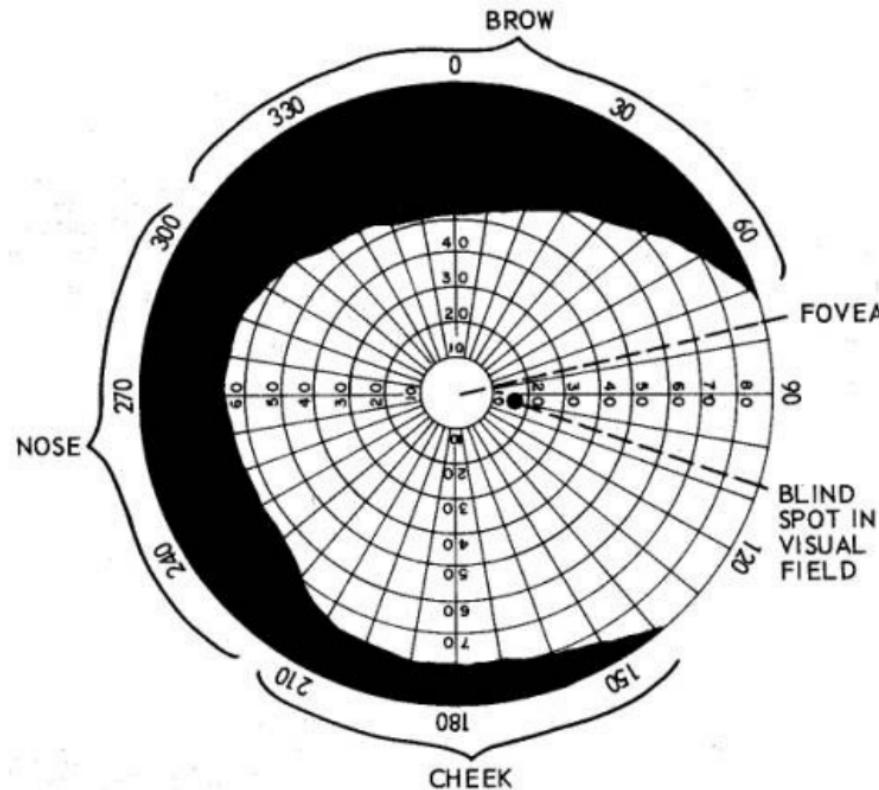


Visual Field

Ruch & Fulton, 1960

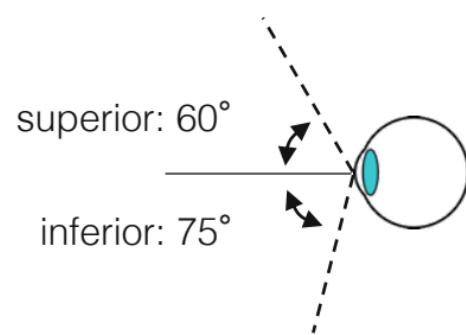


monocular visual field of right eye

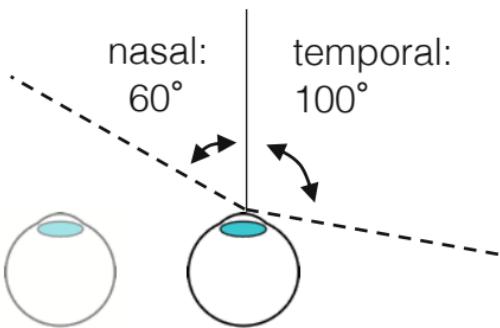


monocular visual field of right eye

Visual Field

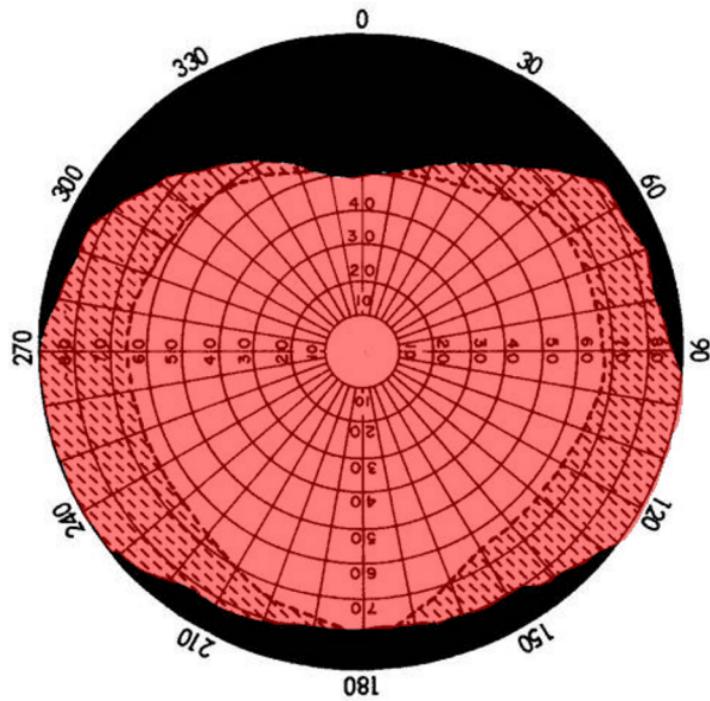


vertical (superior / inferior directions)



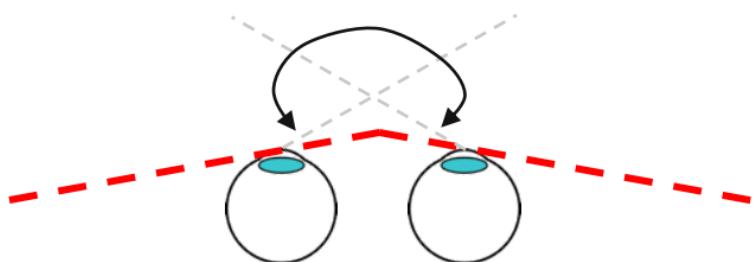
horizontal (nasal / temporal directions)

Visual Field



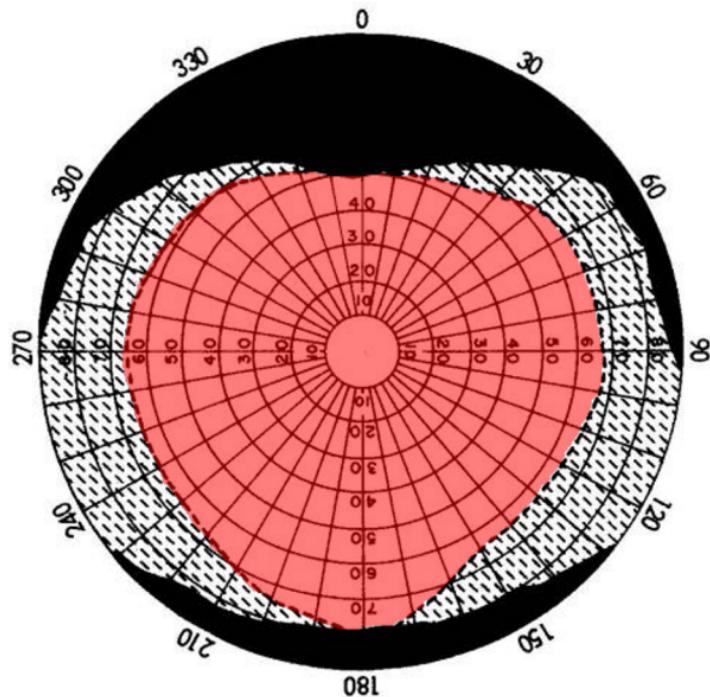
visual field of both eyes

temporal left + temporal right: 200°



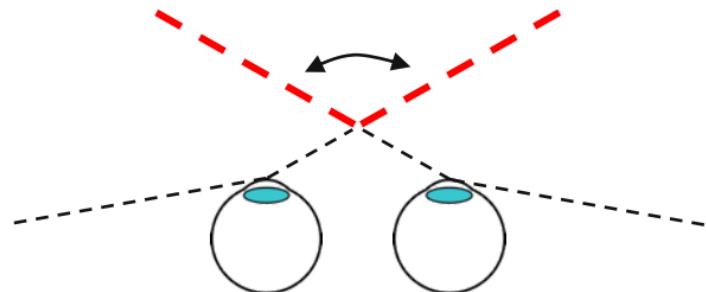
total visual field of both eyes

Visual Field



visual field of both eyes

nasal left + nasal right: 120°

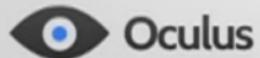
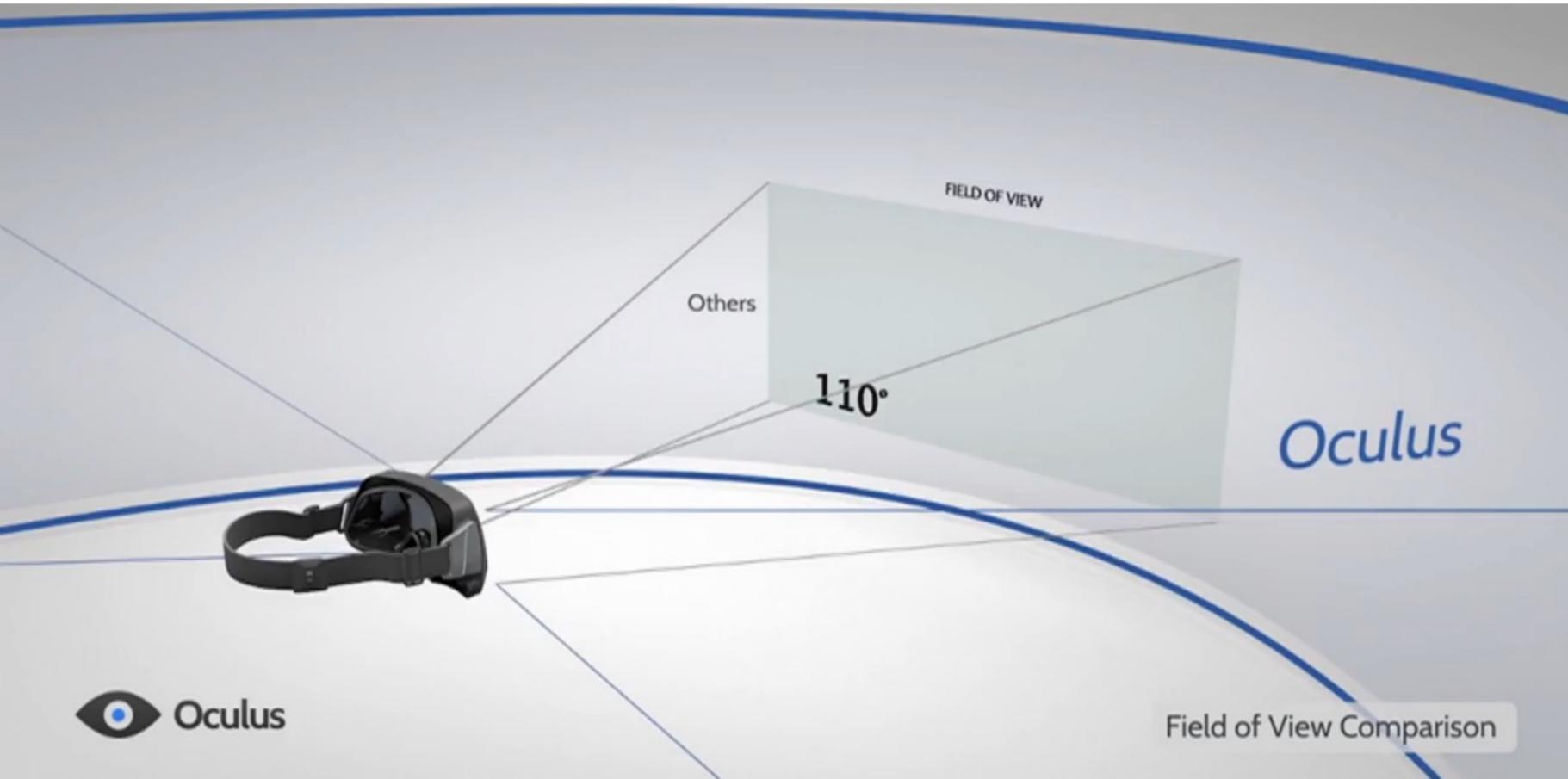


binocular visual field or
region of binocular overlap
→ stereo vision

Visual Field - Terminology

- monocular visual field: visual field of either only the left or right eye
- binocular visual field or region of binocular overlap: intersection of monocular visual fields, i.e. only the overlapping part of both eyes – this is where we see stereo!
- total visual field: union of monocular visual fields, i.e. visual fields of both eyes combined – not all of this is stereo, temporal peripheries are mono!

Immersive VR – How Important is the FOV?

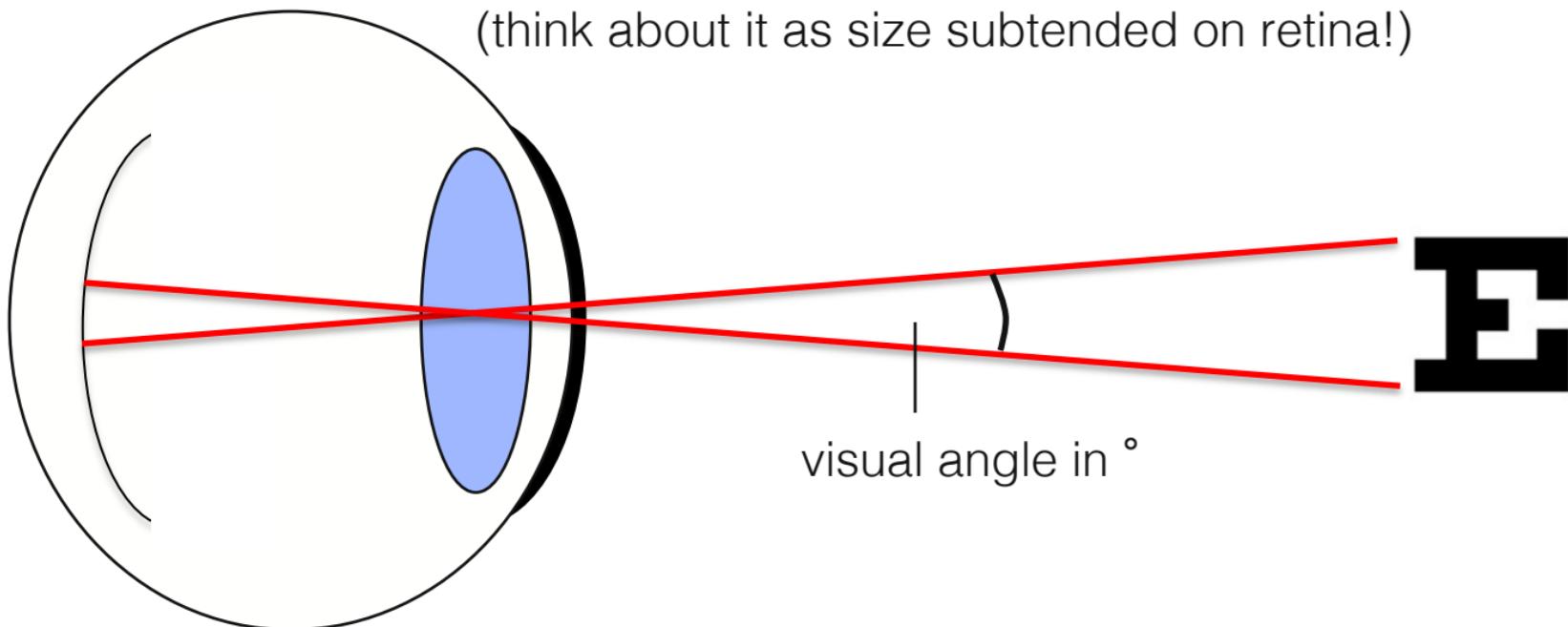


Oculus

Field of View Comparison

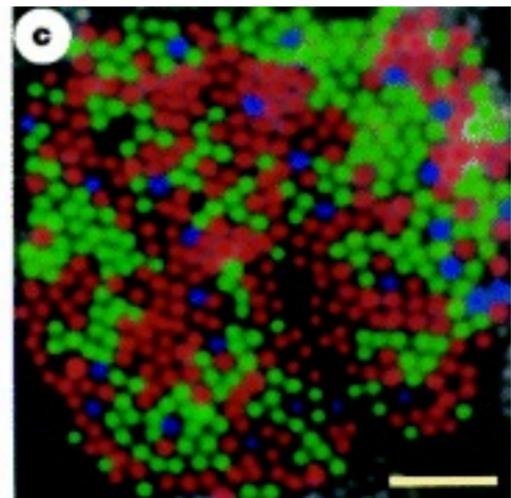
Visual Angle

- vision scientists often measure size in visual angle
- visual angle \approx object size / object distance in degree
(think about it as size subtended on retina!)



Visual Acuity

each photoreceptor
~ 1 arc min (1/60 of a degree)
of visual angle



5 arcmin visual angle

Visual Acuity

Snellen chart



1 20/200

2 20/100

3 20/70

4 20/50

5 20/40

6 20/30

7 20/25

8 20/20

9

10

11



characters are 5 arc min of visual angle, need to resolve 1 arc min to read

Retina VR Display – What does it Take?

need per eye:

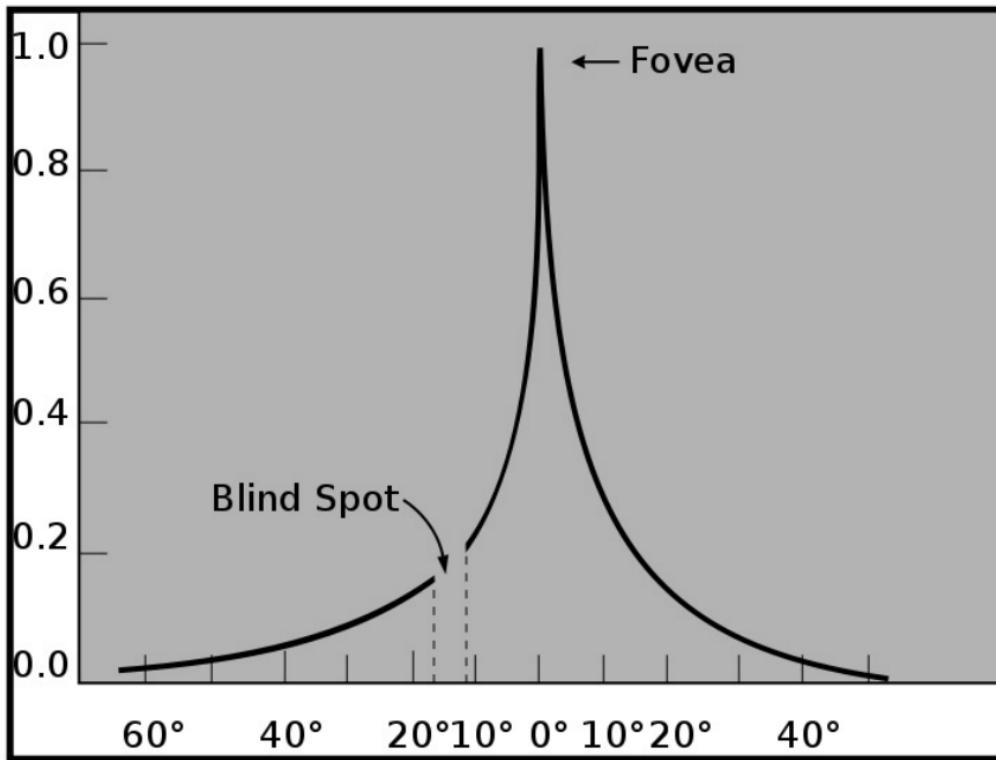
150° x 135° with pixels covering 1 arc min of visual angle
= 9000 x 8100 pixels (probably 2-3x of that in practice)

biggest challenge: bandwidth

- capture or render stereo panoramas or images at that resolution
- compress and transmit huge amount of data
- drive and operate display pixels

Relative Acuity Over Retina

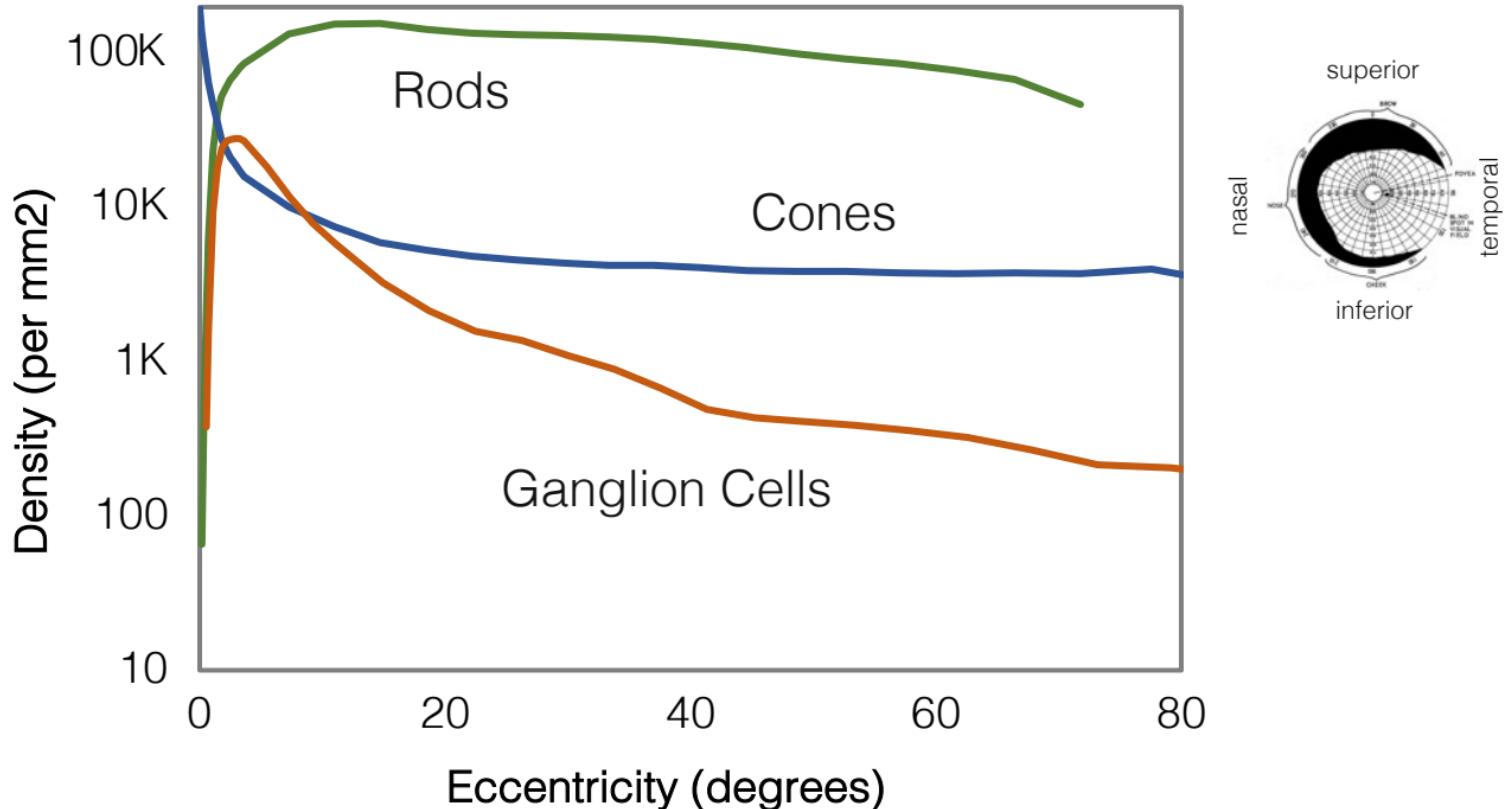
wikipedia



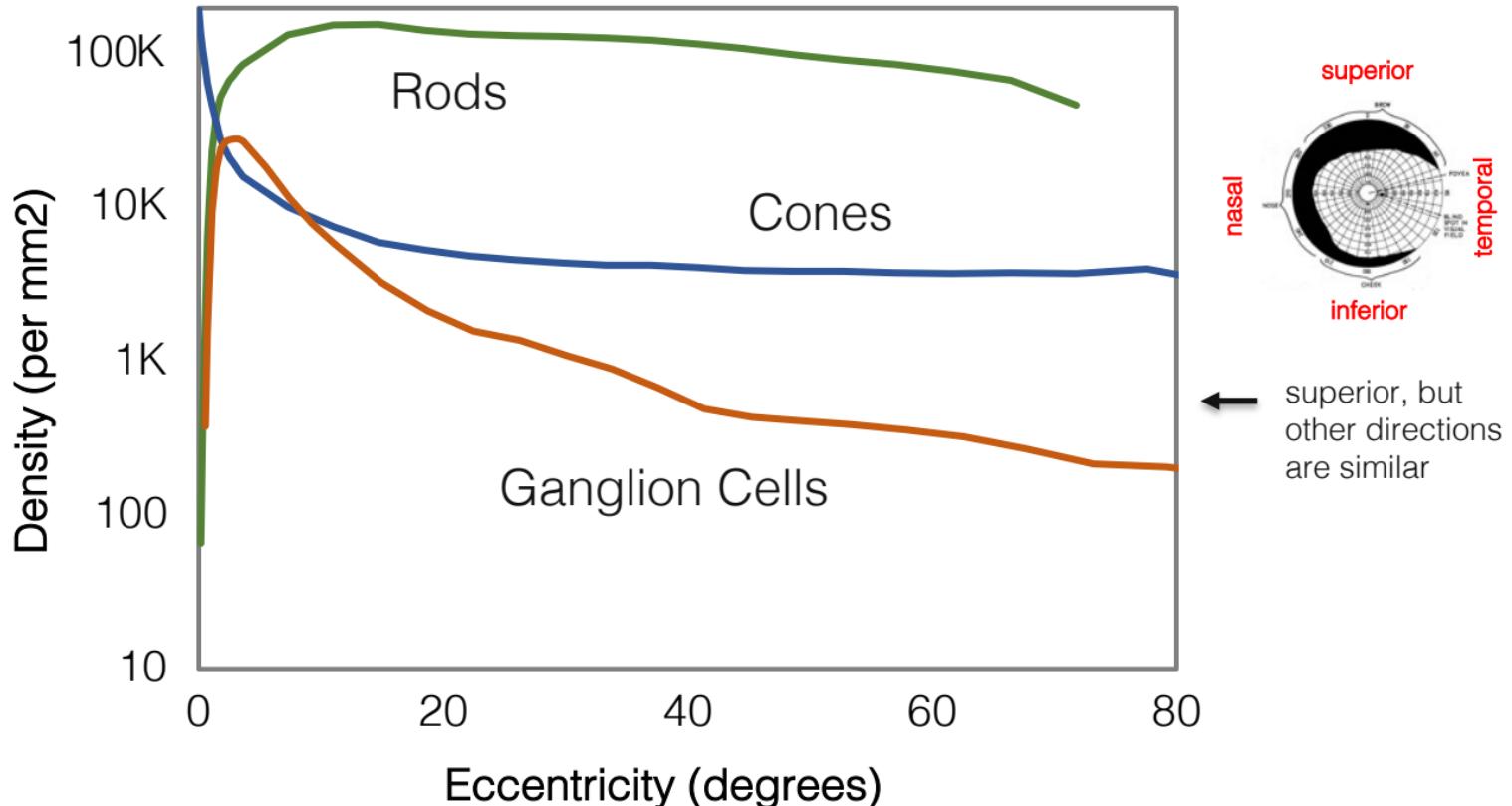
Eccentricity (i.e., distance to fovea in degrees of visual angle)

Density of Photoreceptors on Retina

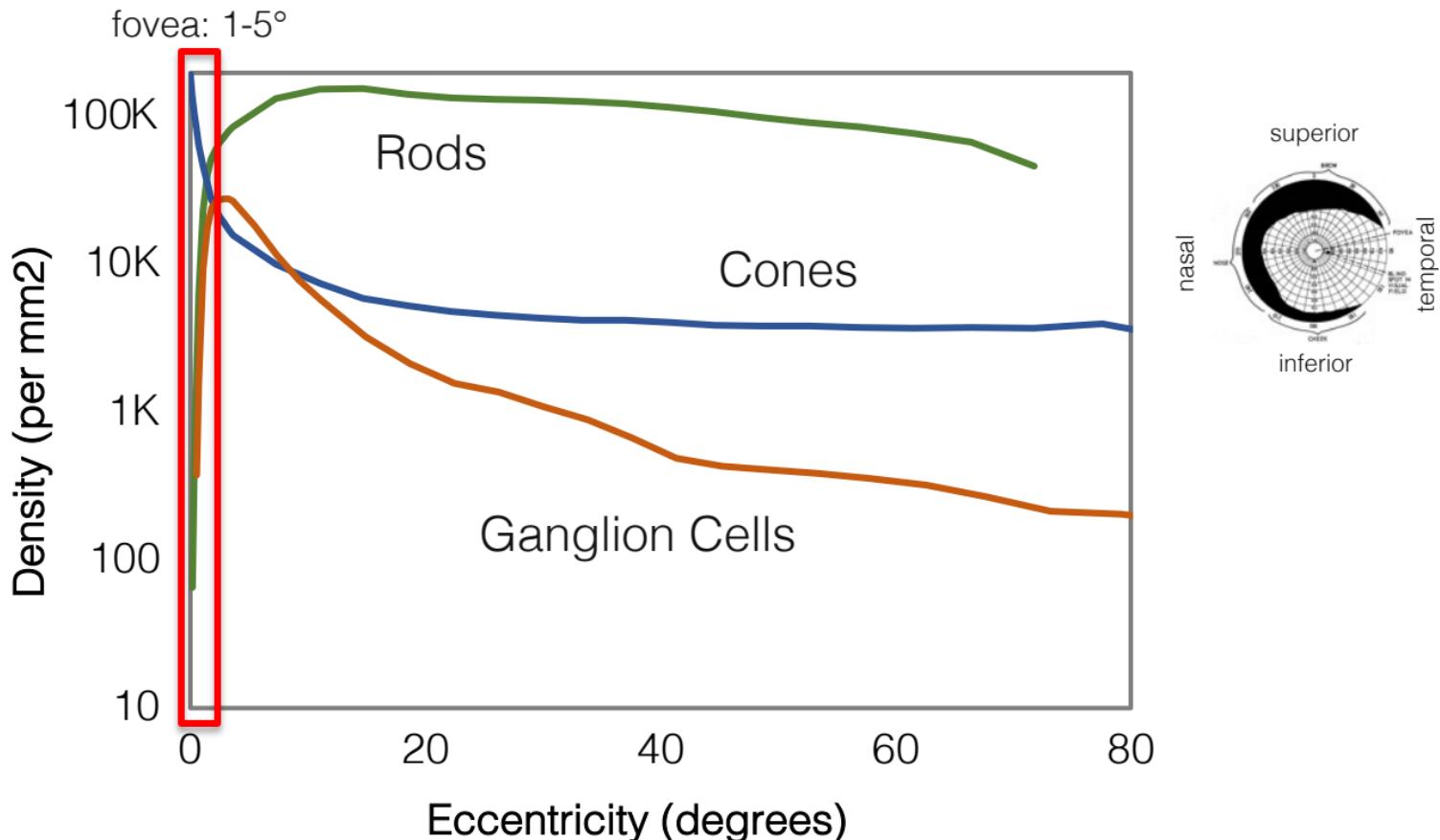
Patney et al. 2016



Density of Photoreceptors on Retina



Density of Photoreceptors on Retina



Acuity Over Retina / MAR

acuity falls off due to:

- reduced receptor and ganglion cell density
- reduced optical nerve “bandwidth”
- reduced “processing” devoted to periphery in the visual cortex

Acuity Over Retina / MAR

acuity falls off due to:

- reduced receptor and ganglion cell density
- reduced optical nerve “bandwidth”
- reduced “processing” devoted to periphery in the visual cortex

MAR: minimum angle of
resolution in deg/cycle slope

$$\omega = me + \omega_0$$

↑ ↑
eccentricity in degrees smallest resolvable angle at fovea in deg/cycle

Acuity Over Retina / MAR

acuity falls off due to:

- reduced receptor and ganglion cell density
- reduced optical nerve “bandwidth”
- reduced “processing” devoted to periphery in the visual cortex

MAR: minimum angle of
resolution in deg/cycle slope

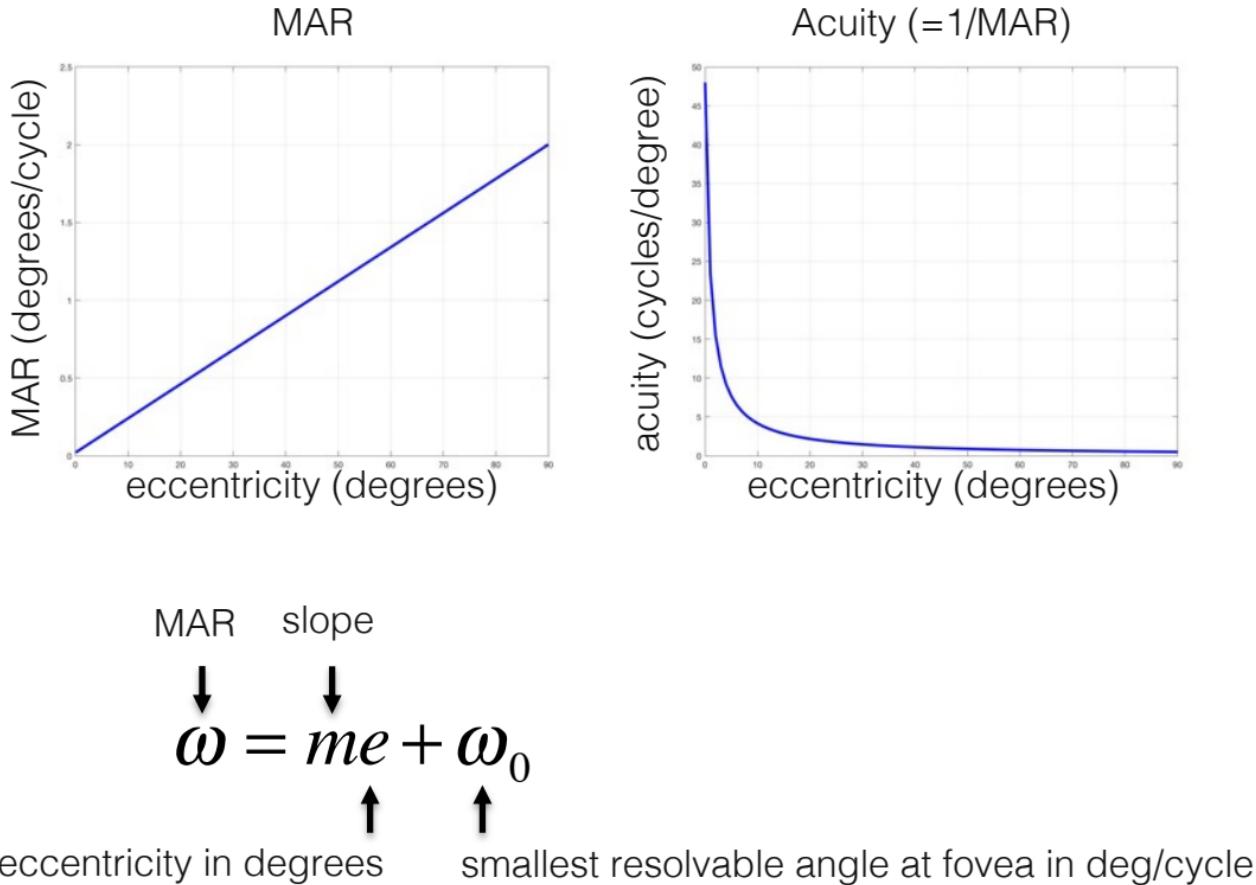
$$\omega_0 = (1/48)^\circ \quad \text{somewhere between 20/20 (30 cycles per degree) and 20/10 (60 cycles per degree)}$$

$$\omega = me + \omega_0$$

↓ ↓
eccentricity in degrees smallest resolvable angle at fovea in deg/cycle

$$m = 0.022 - 0.034 \quad \text{range of acceptable – equivalent for observed image quality}$$

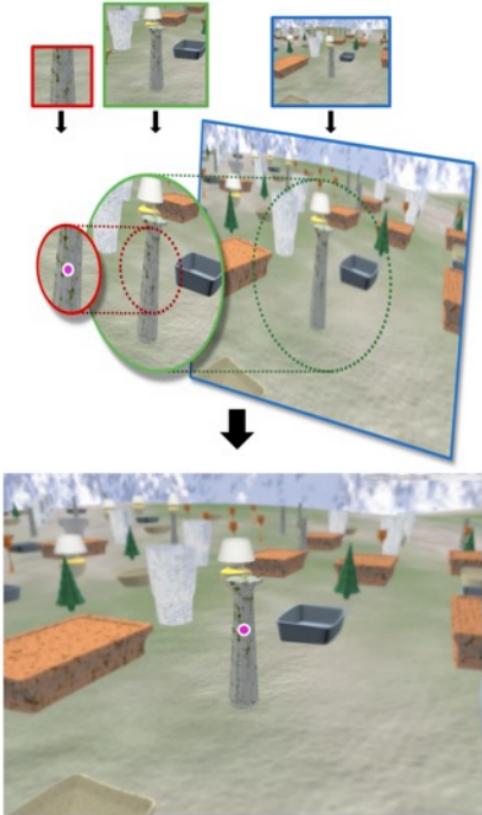
Acuity Over Retina / MAR







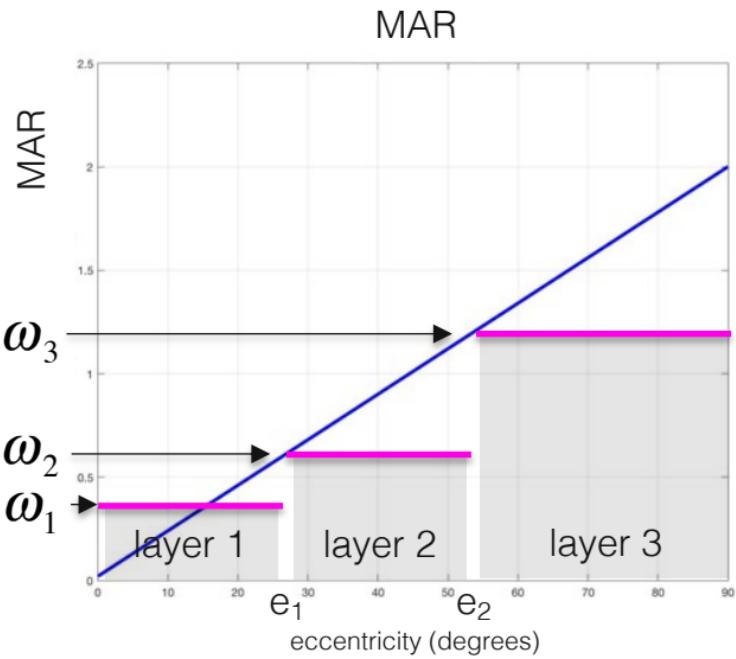
Foveated Rendering



- Guenter et al. 2012: split image into n layers, e.g. inner (foveal, 1), middle (2), outer (3)
- render image in each zone with progressively lower resolution
- goals: save computation & bandwidth!

Foveated Rendering

- Guenter et al. 2012: split image into n layers, e.g. inner (foveal, 1), middle (2), outer (3)



$$e_i = \frac{i}{n} \cdot \frac{fov}{2}$$

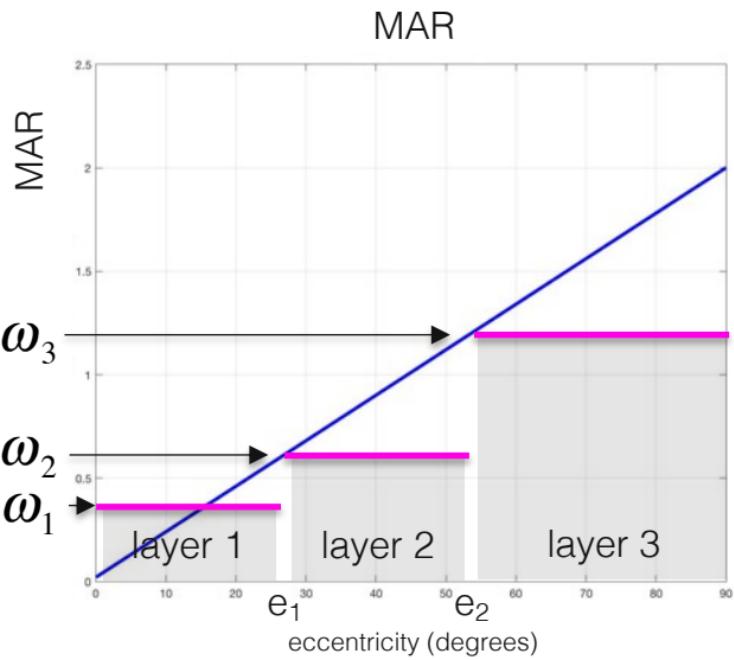
➡

$$e_1 = \frac{fov}{6}$$
$$e_2 = \frac{fov}{3}$$

Foveated Rendering

ω_1 is best the display can do!

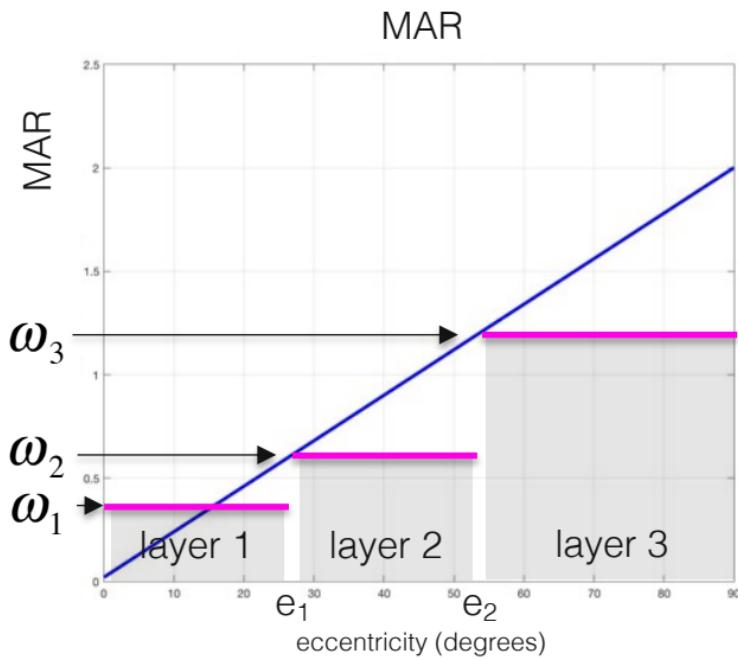
$$\text{unit of } \omega_1: \frac{\text{degrees}}{\text{cycle}} = \frac{\text{degrees}}{2 \cdot \text{pixel_size}}$$



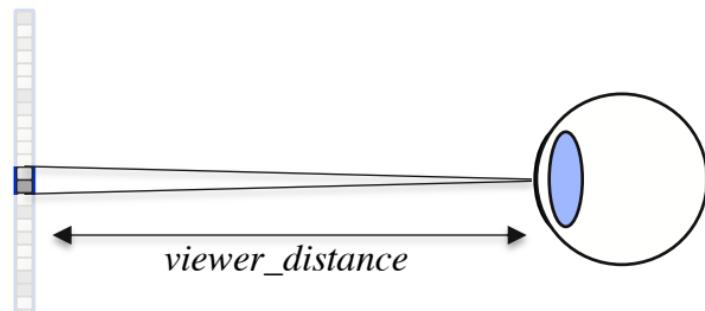
Foveated Rendering

ω_1 is best the display can do!

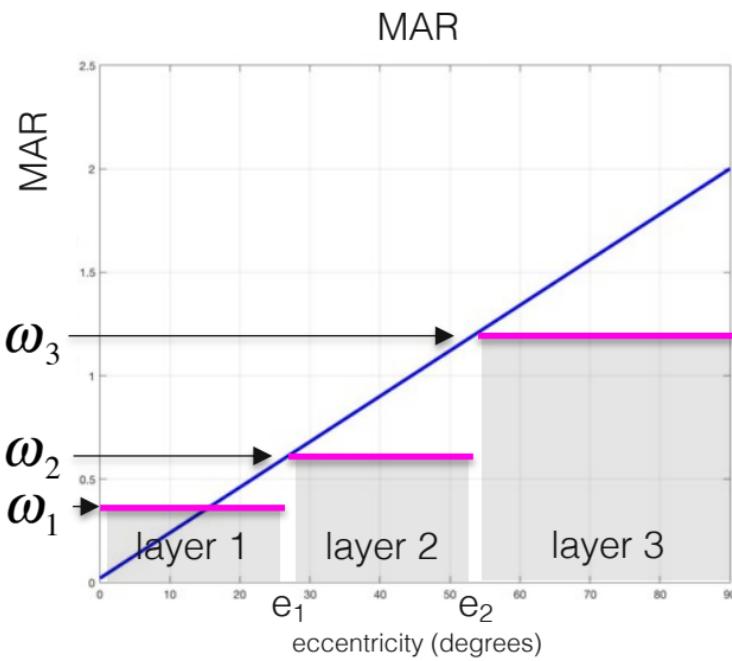
unit of ω_1 : $\frac{\text{degrees}}{\text{cycle}} = \frac{\text{degrees}}{2 \cdot \text{pixel_size}}$



$$\omega_1 = 2 \tan^{-1} \left(\frac{\text{screen_size}}{\text{screen_resolution} \cdot \text{viewer_distance}} \right) \cdot \frac{360}{2\pi}$$



Foveated Rendering



ω_1 is best the display can do!

unit of ω_1 : $\frac{\text{degrees}}{\text{cycle}} = \frac{\text{degrees}}{2 \cdot \text{pixel_size}}$

$$\omega_1 = 2 \tan^{-1} \left(\frac{\text{screen_size}}{\text{screen_resolution} \cdot \text{viewer_distance}} \right) \cdot \frac{360}{2\pi}$$

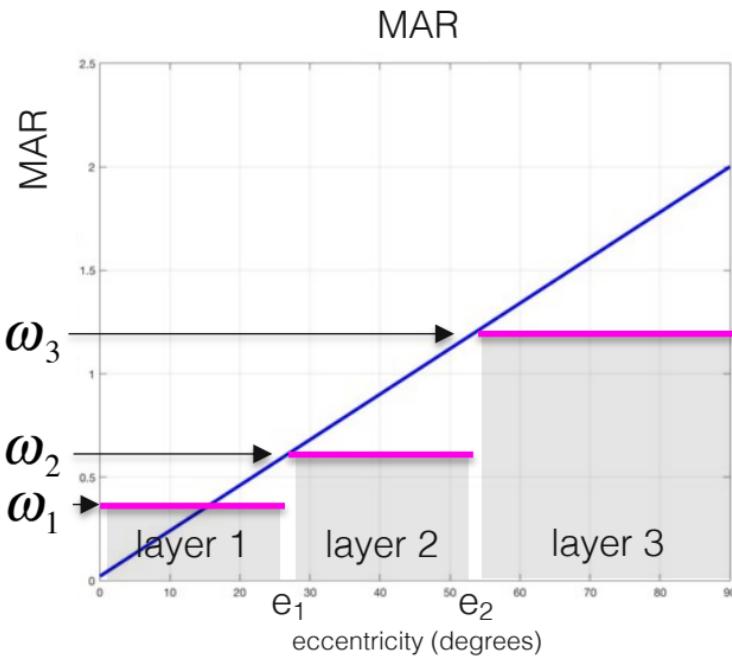
screen_size

is *either* screen width *or* height (same units as viewer_distance)

screen_resolution

is *either* number of horizontal pixels *or* vertical pixels of the screen (same dimension as screen_size)

Foveated Rendering



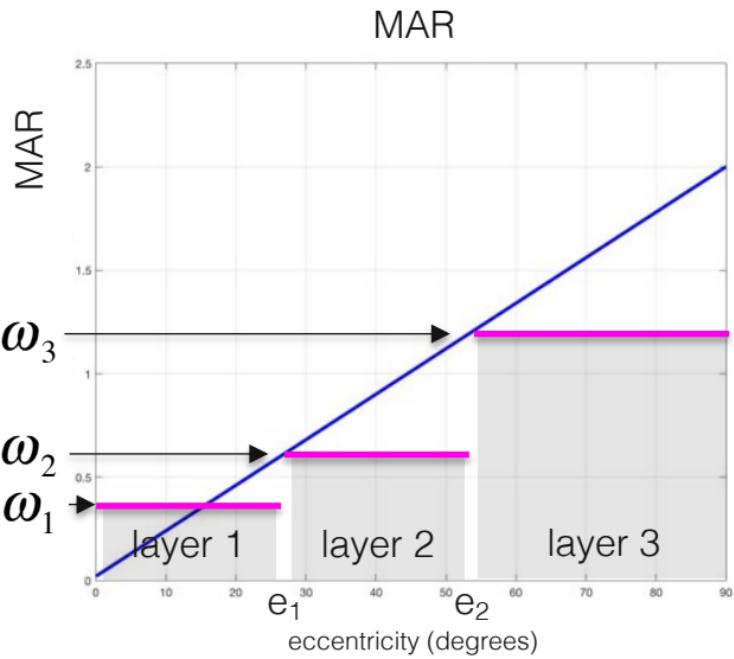
$$\omega_1 = 2 \tan^{-1} \left(\frac{\text{screen_size}}{\text{screen_resolution} \cdot \text{viewer_distance}} \right) \cdot \frac{360}{2\pi}$$

$$\omega_2 = m e_2 + \omega_0$$

$$\omega_3 = m e_3 + \omega_0$$

Foveated Rendering

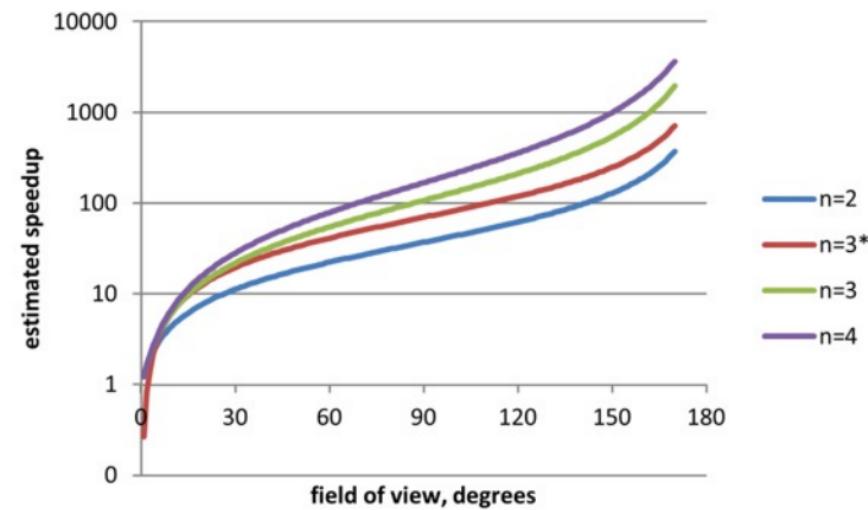
- convert MAR (in degrees/cycle) to pixels



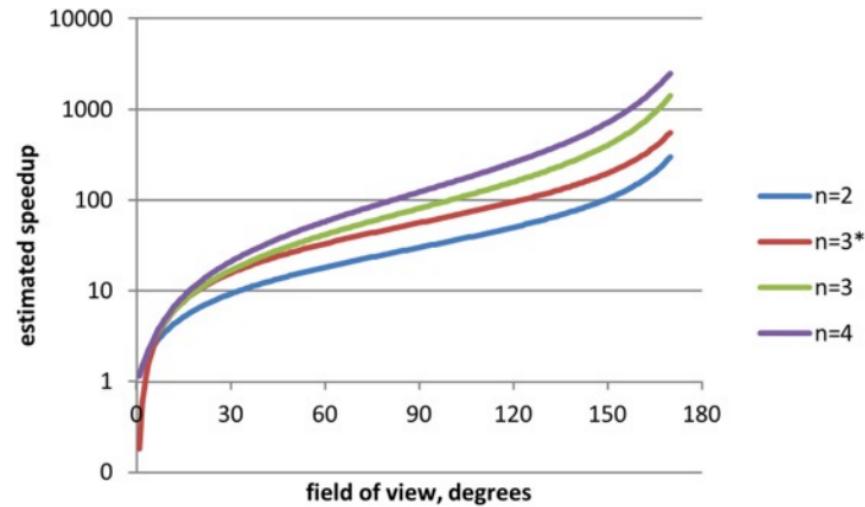
$$\text{blur_radius_in_px} = \text{viewer_distance} \cdot \tan\left(\frac{\omega}{2} \cdot \frac{2\pi}{360}\right)$$

Foveated Rendering – Performance Gain

$$m = 0.028$$



$$m = 0.022$$



n is number of layers

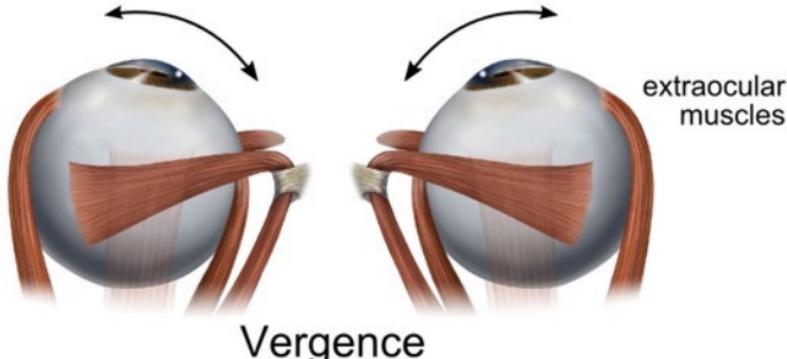
speedup is total number of display pixels / number of pixels in all layers combined

conclusion: for large fov & high-res displays, we need to shade much fewer pixels!

Depth Perception

Oculomotor Cue

Stereopsis (Binocular)

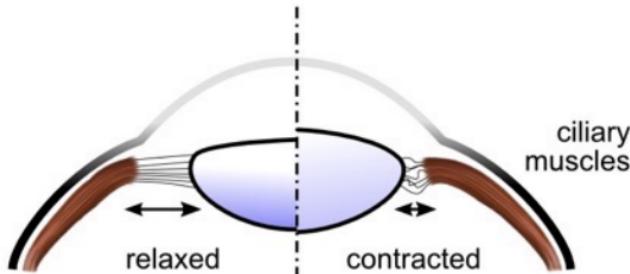


Vergence



Binocular Disparity

Focus Cues (Monocular)



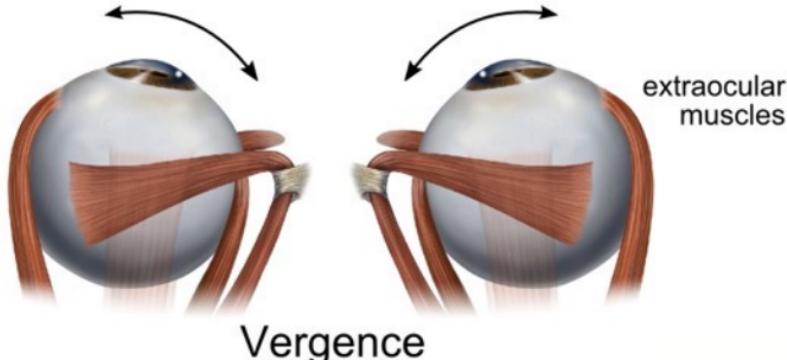
Accommodation



Retinal Blur

Oculomotor Cue

Stereopsis (Binocular)



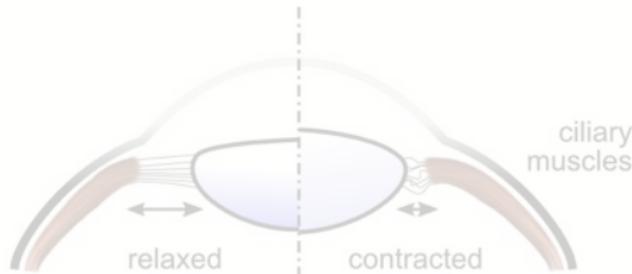
Vergence



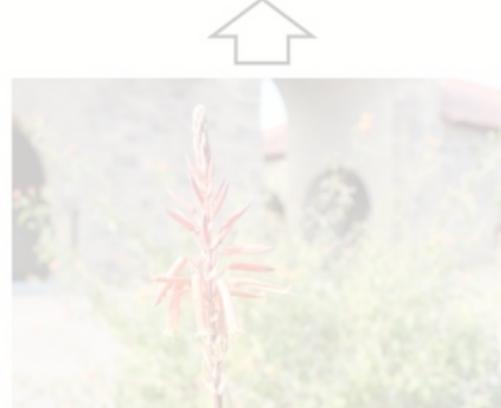
Binocular Disparity

Visual Cue

Focus Cues (Monocular)



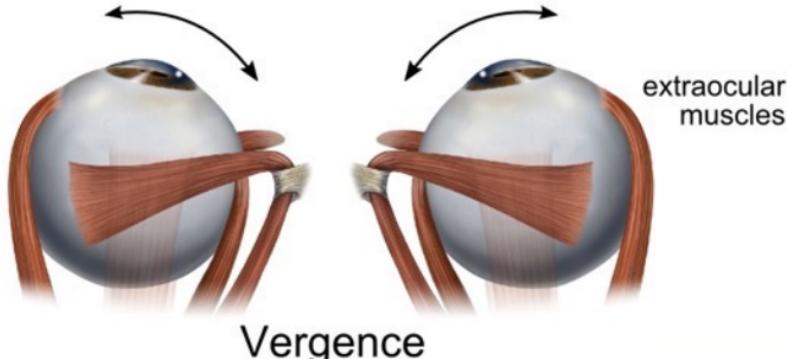
Accommodation



Retinal Blur

Oculomotor Cue

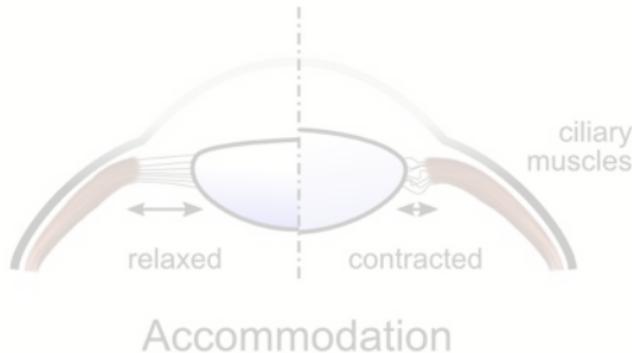
Stereopsis (Binocular)



Binocular Disparity

Visual Cue

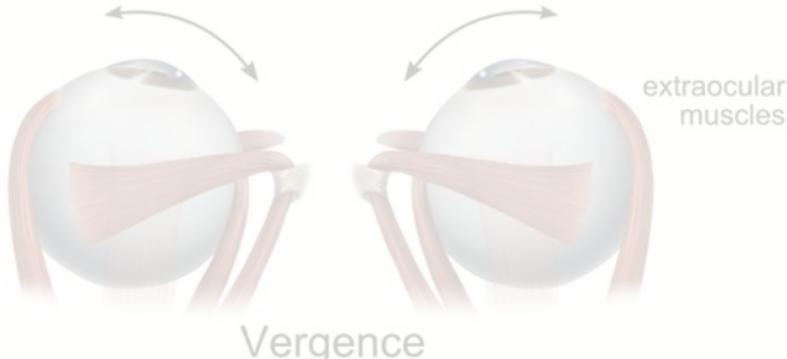
Focus Cues (Monocular)



Retinal Blur

Oculomotor Cue

Stereopsis (Binocular)

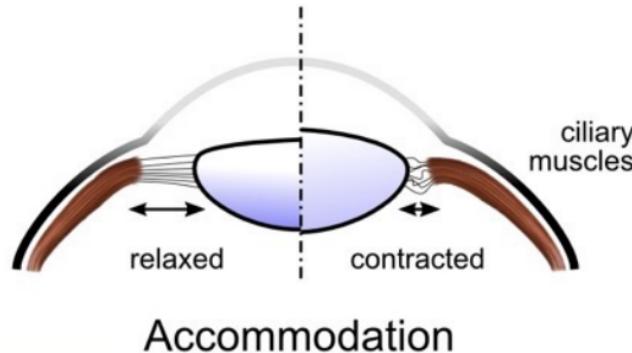


Visual Cue

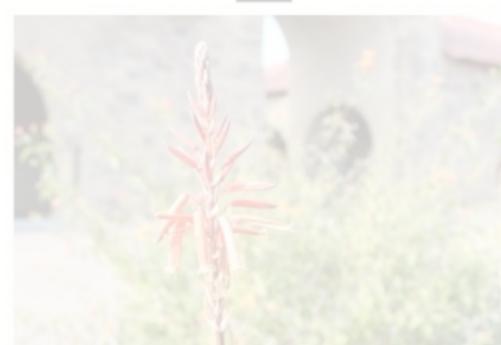
Binocular Disparity



Focus Cues (Monocular)



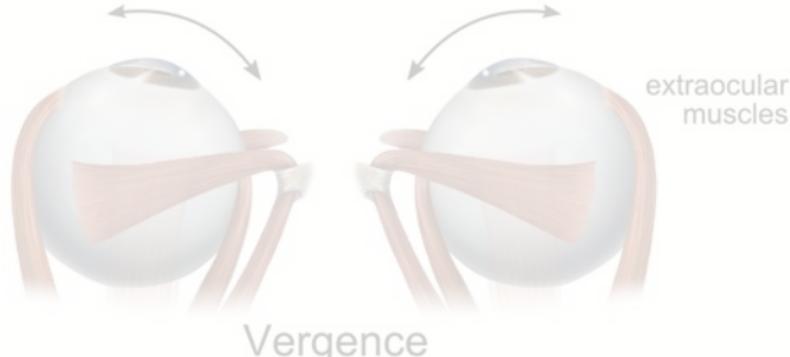
Accommodation



Retinal Blur

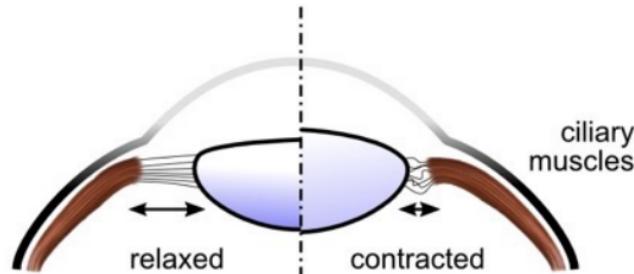
Oculomotor Cue

Stereopsis (Binocular)



Binocular Disparity

Focus Cues (Monocular)



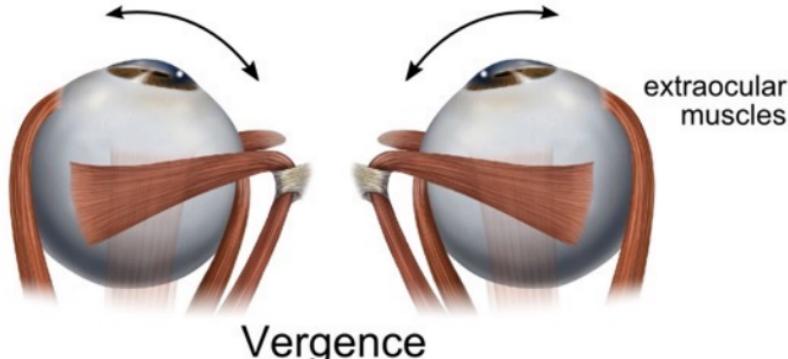
Accommodation



Retinal Blur

Oculomotor Cue

Stereopsis (Binocular)

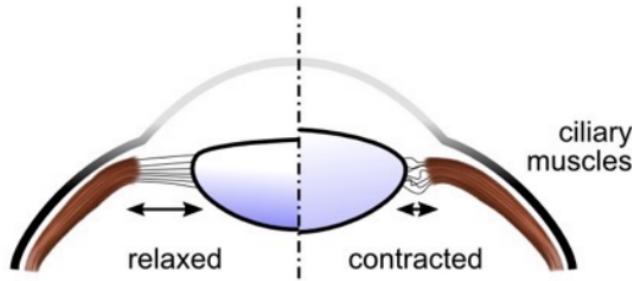


Vergence



Binocular Disparity

Focus Cues (Monocular)



Accommodation



Retinal Blur

Visual Cue

Depth Perception



Depth Perception



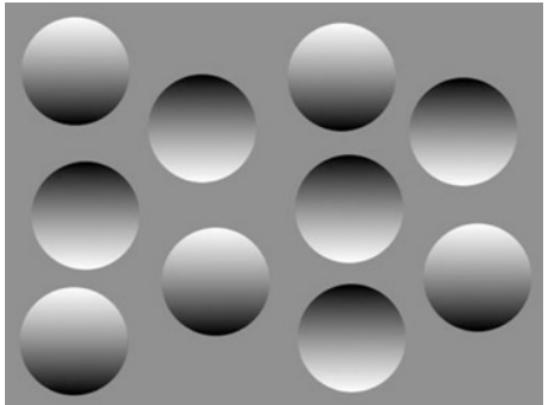
monocular cues

- perspective
- relative object size
- absolute size
- occlusion
- accommodation
- retinal blur
- motion parallax
- texture gradients
- shading
- ...

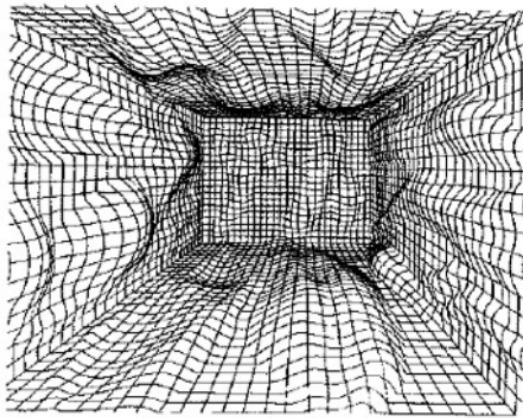
binocular cues

- (con)vergence
- disparity / parallax
- ...

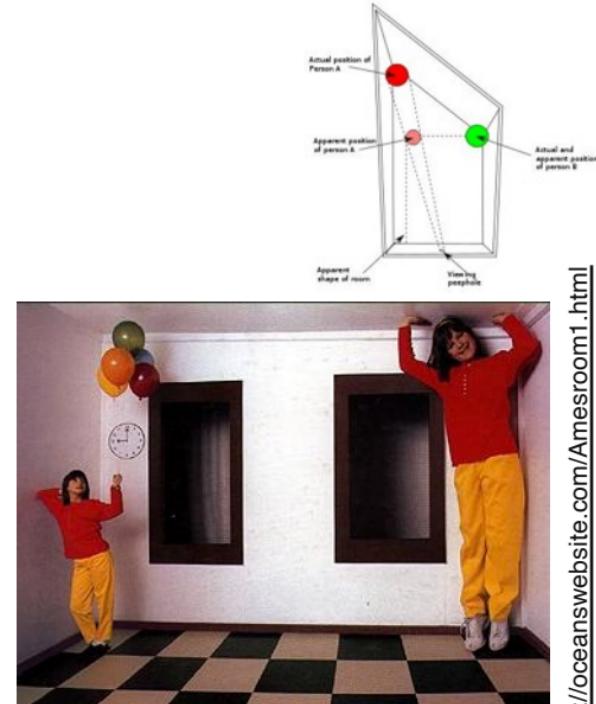
Depth Perception – Pictorial Cues



shading

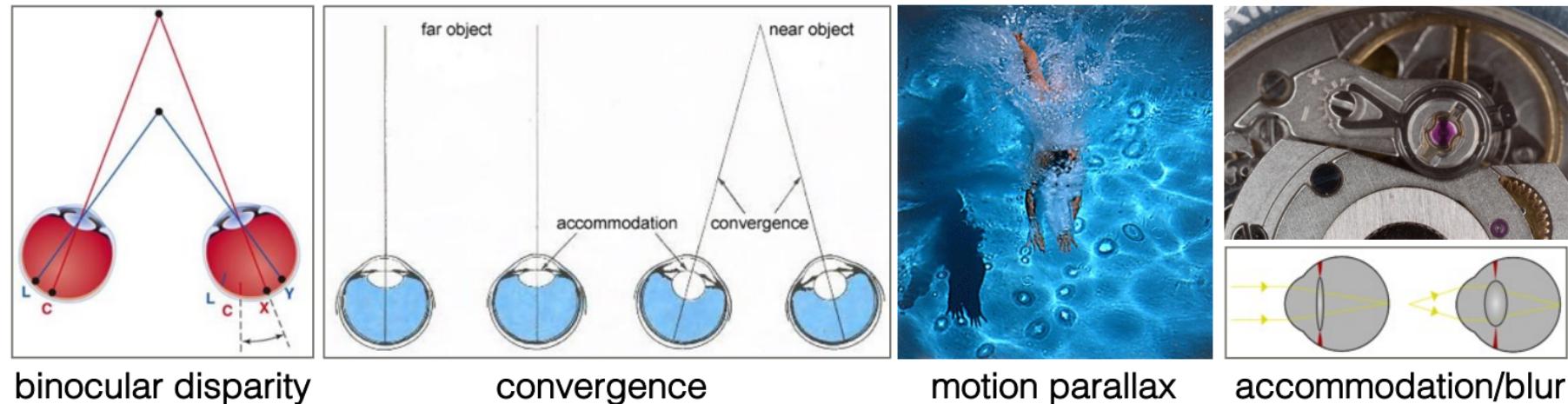


texture gradients



Ames room
conflict between perspective
& relative object size

Depth Perception



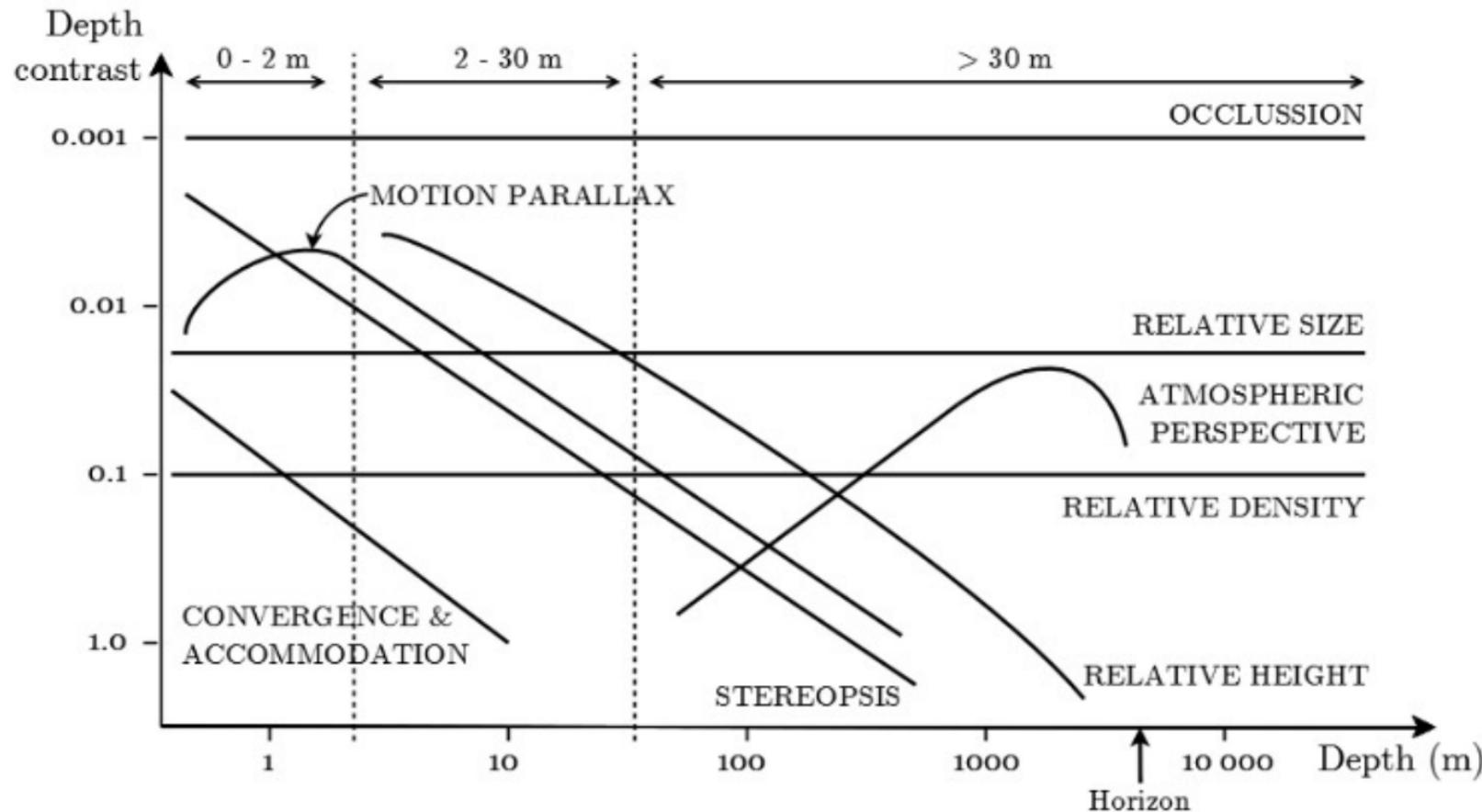
↔ current glasses-based (stereoscopic) displays

↔ near-term: light field displays

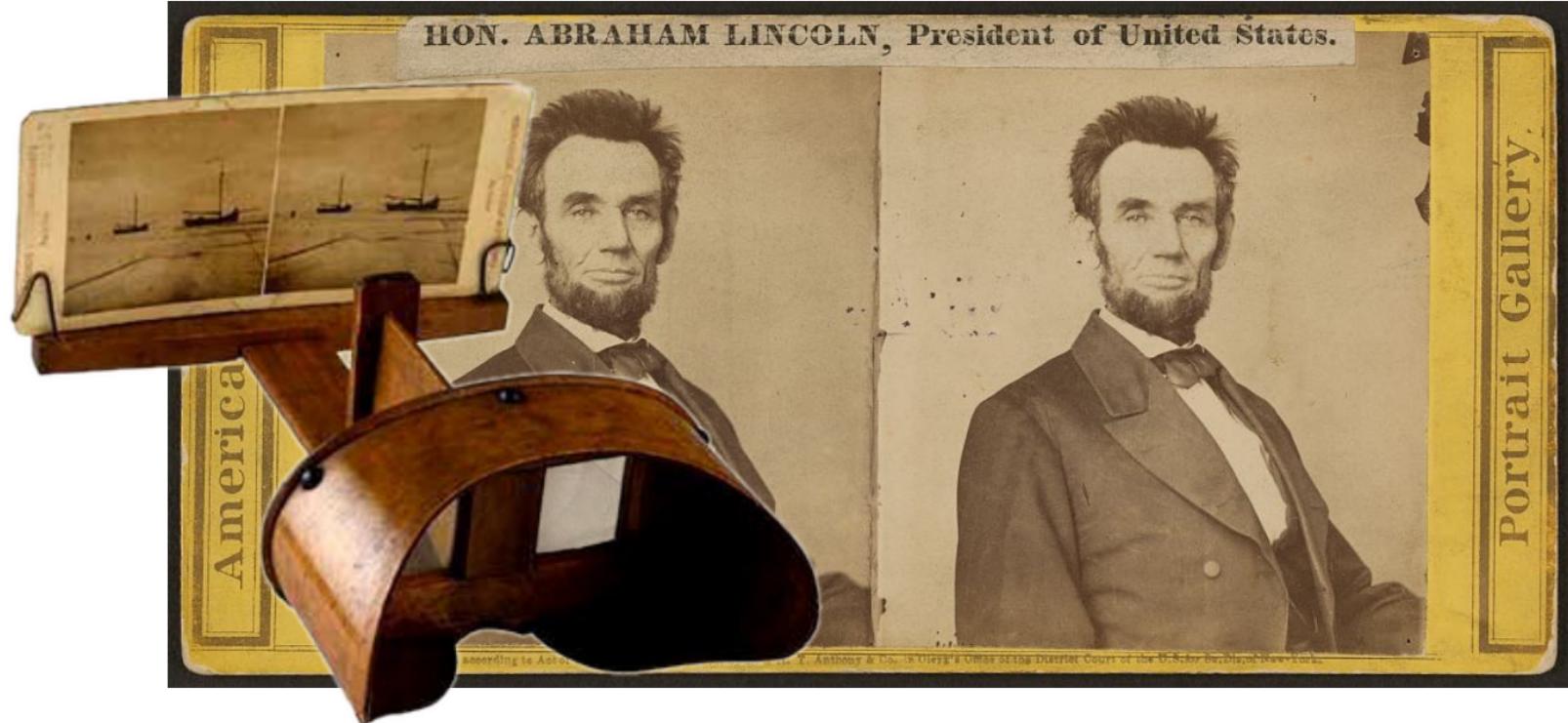
↔ longer-term: holographic displays

Depth Perception

Cutting & Vishton, 1995



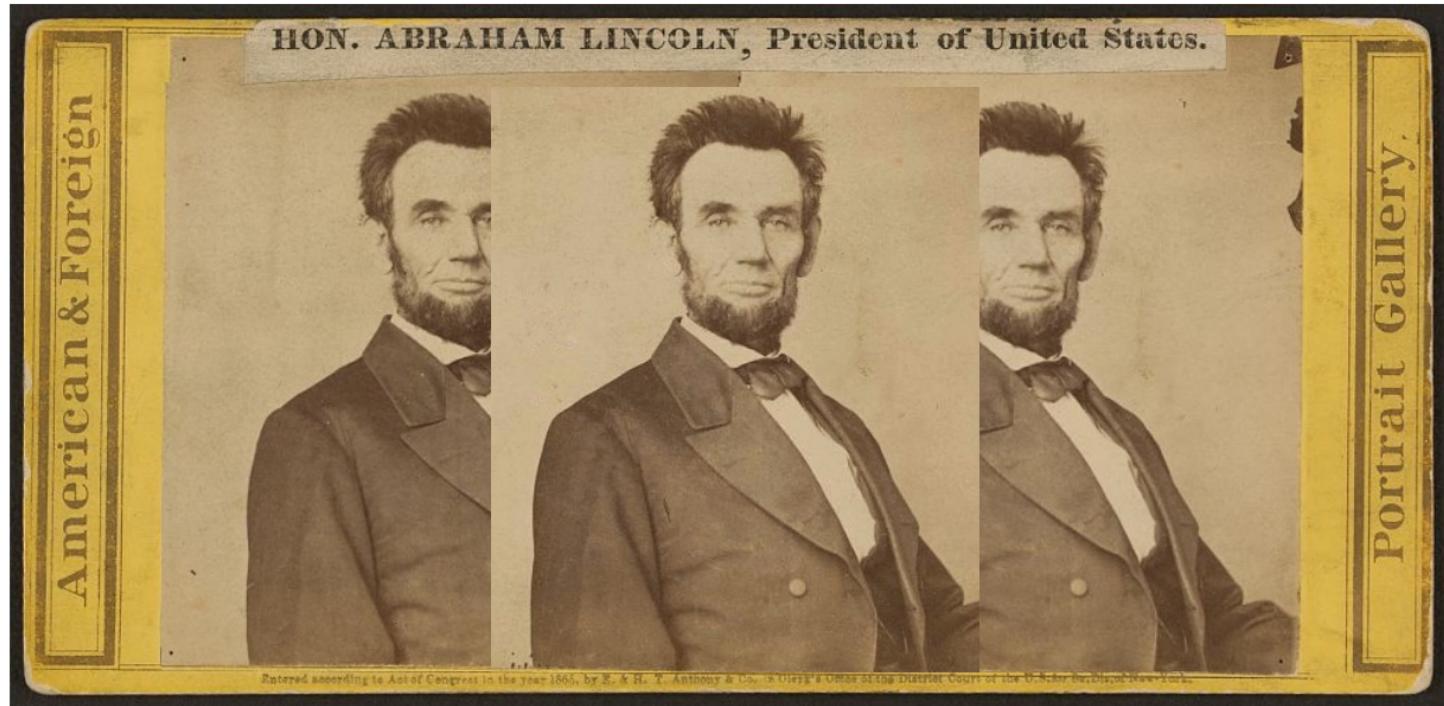
Stereoscopic Displays



Charles Wheatstone., 1841. Stereoscope.

Walker, Lewis E., 1865. Hon. Abraham Lincoln, President of the United States. Library of Congress

Stereoscopic Displays



Stereoscopic Displays



176 years later



Charles Wheatstone 1838

stereoscopic displays

A Brief History of Virtual Reality

Stereoscopes

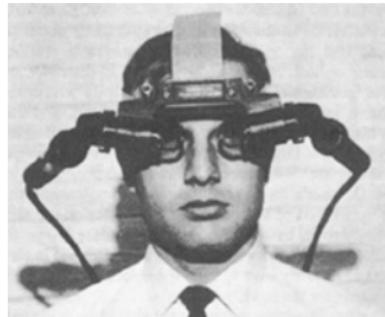
Wheatstone, Brewster, ...



1838

VR, AR,

Ivan Sutherland



1968

VR explosion

Oculus, Sony, Valve, MS, ...



2012-2021

Next-generation VR & AR Displays

Focus Cues

Oculomotor Processes

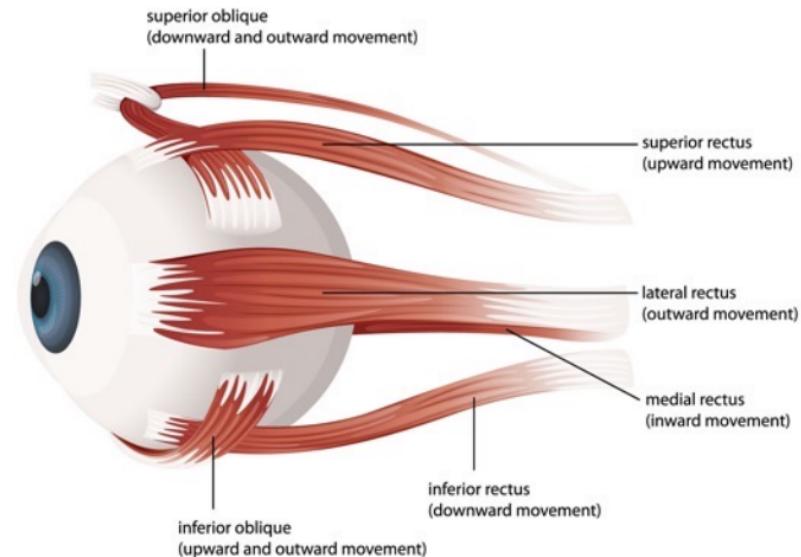
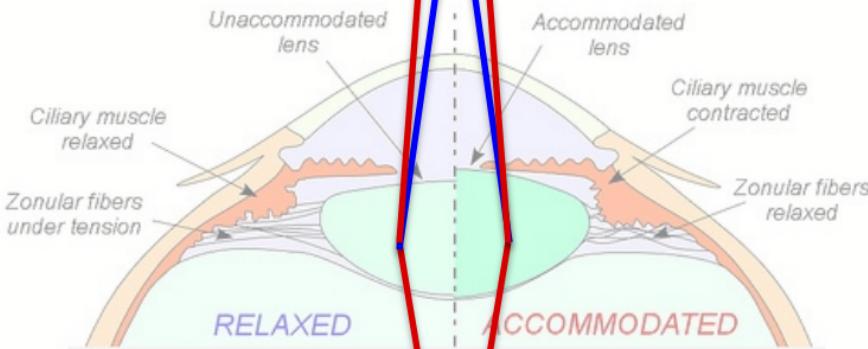
far focus



16 years: ~8cm to ∞

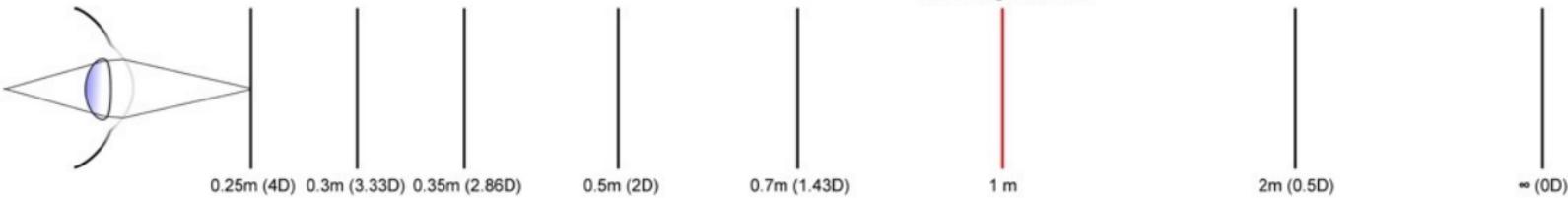
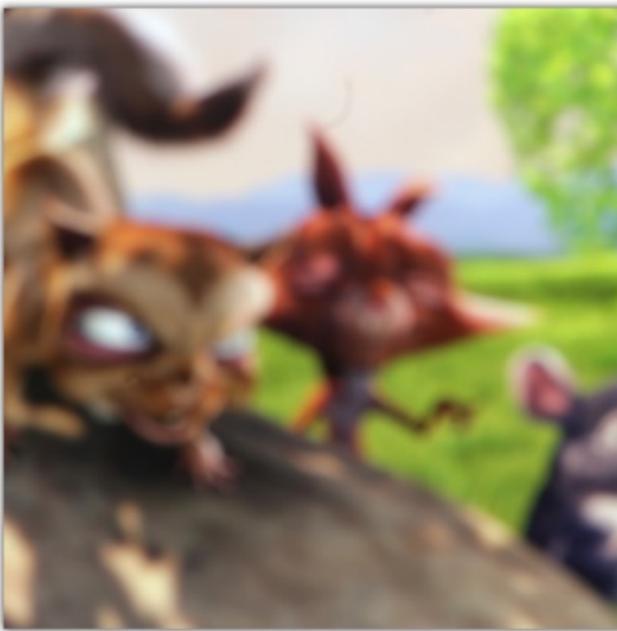
50 years: ~50cm to ∞ (mostly irrelevant)

near focus



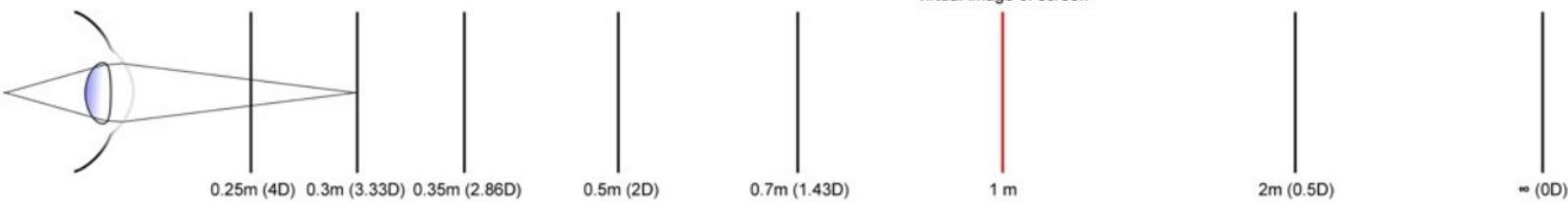
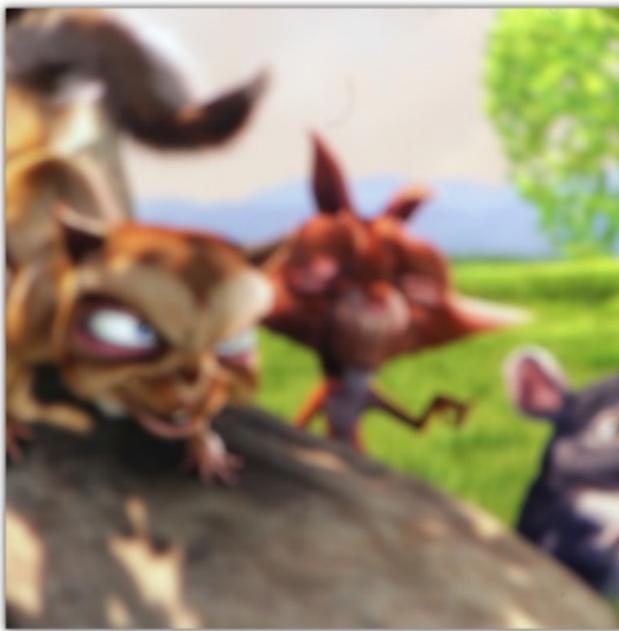
Accommodation and Retinal Blur

Conventional Display



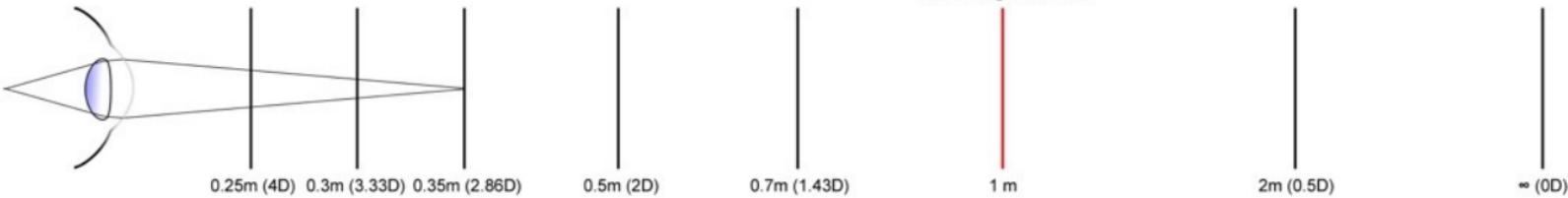
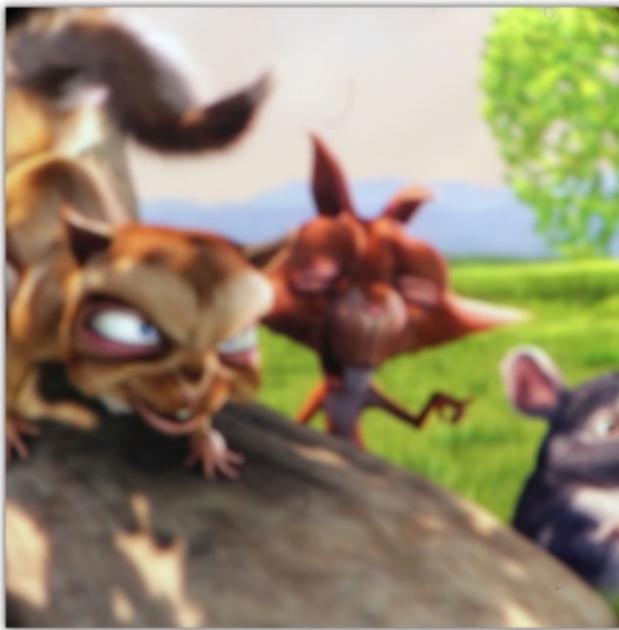
Accommodation and Retinal Blur

Conventional Display



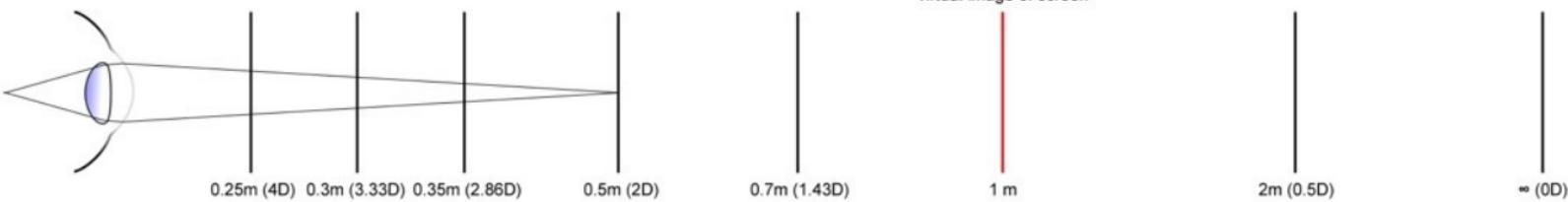
Accommodation and Retinal Blur

Conventional Display



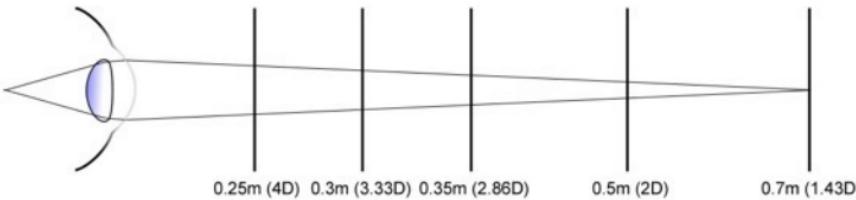
Accommodation and Retinal Blur

Conventional Display



Accommodation and Retinal Blur

Conventional Display



virtual image of screen

1 m

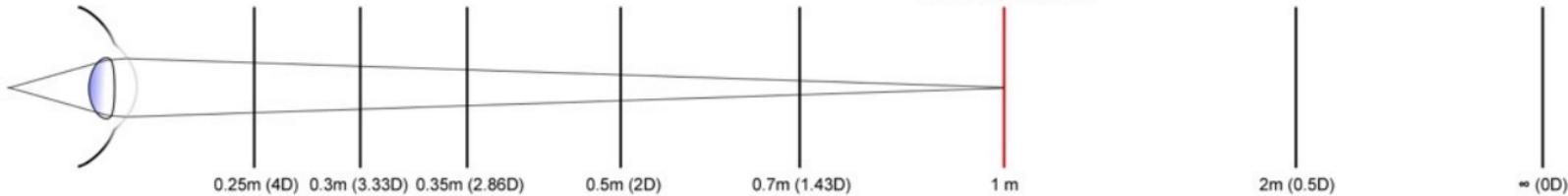


Accommodation and Retinal Blur

Conventional Display

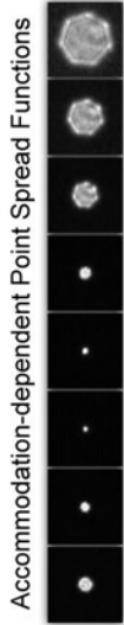


virtual image of screen

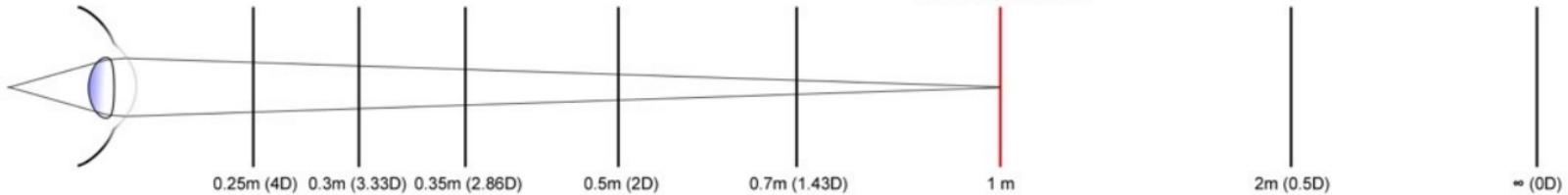


Accommodation and Retinal Blur

Conventional Display

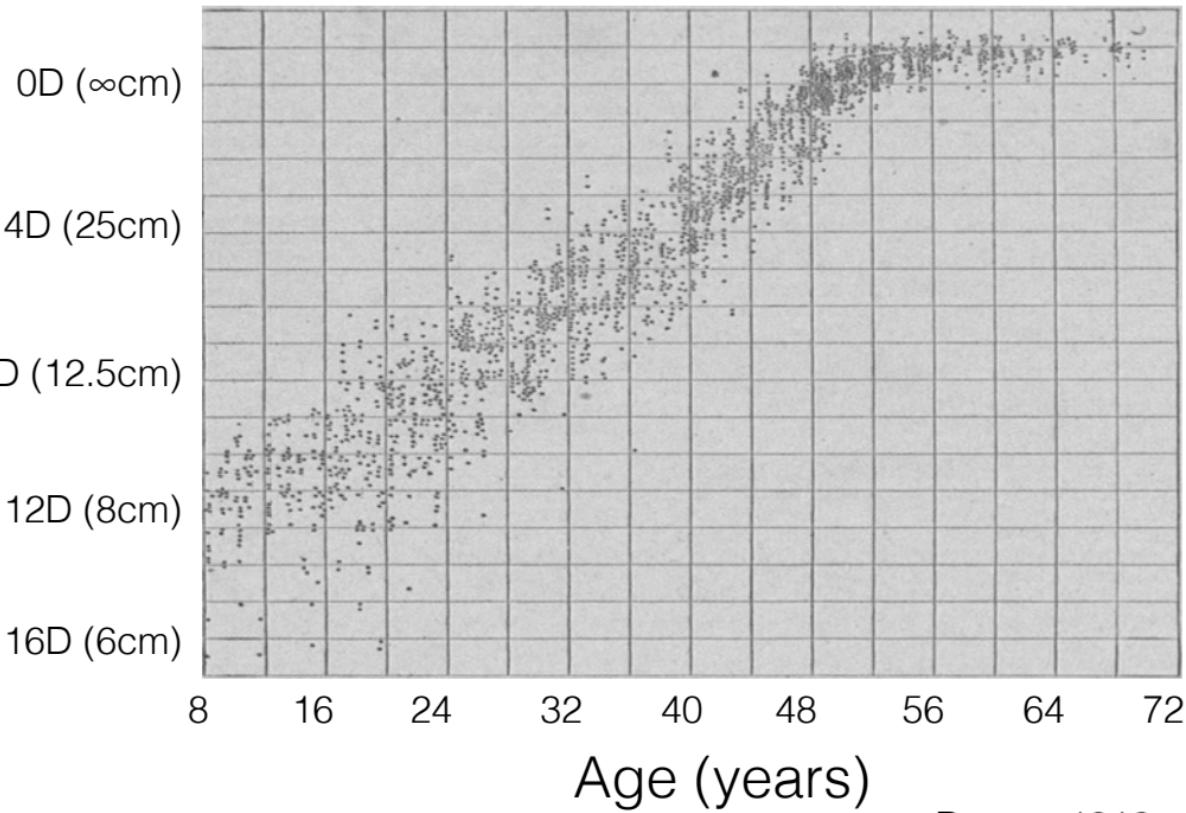


virtual image of screen



Focusing Ability Degrades With Age - Presbyopia

Nearest focus distance

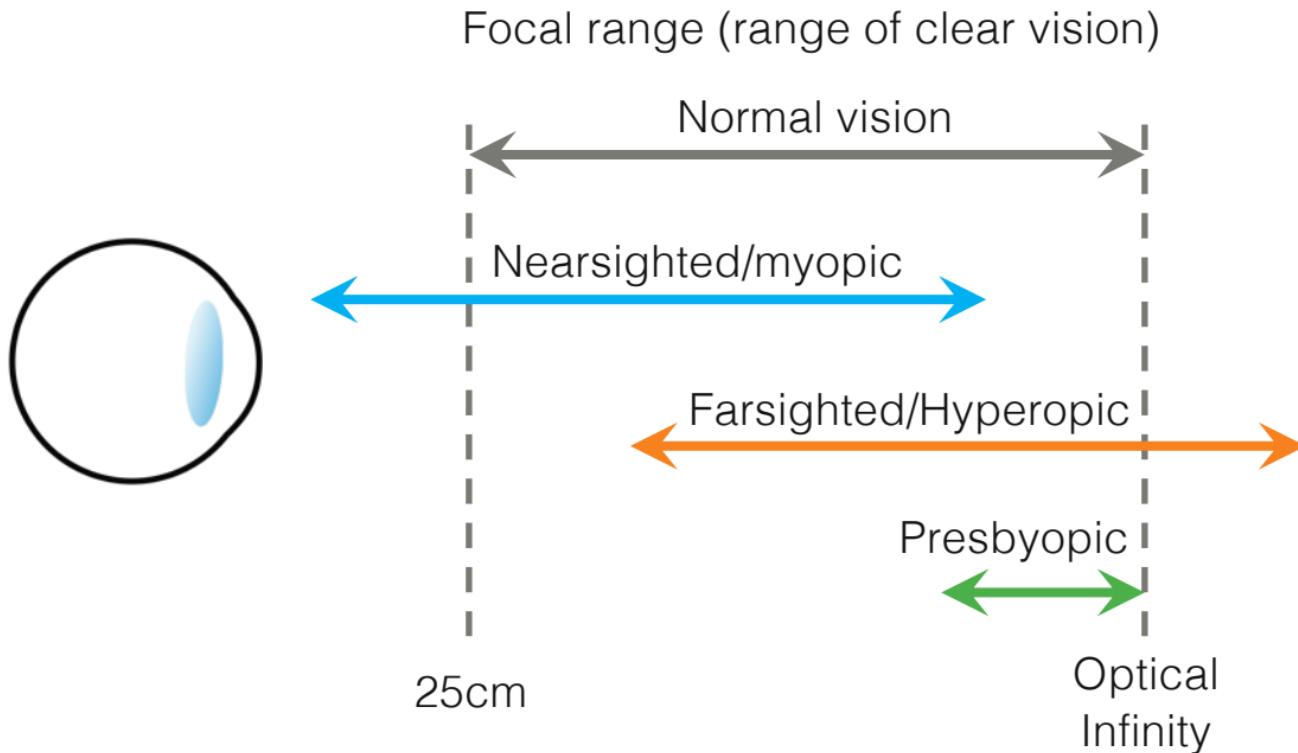


Duane, 1912

Bifocals

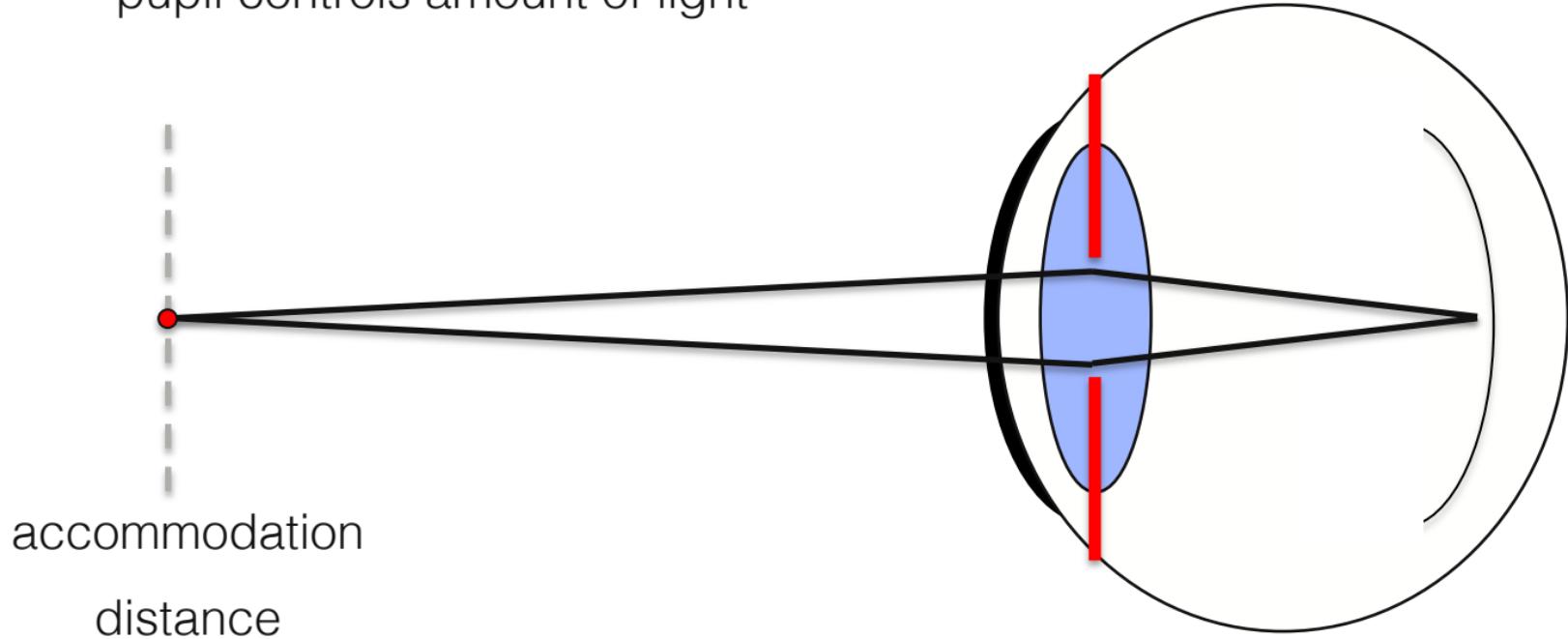


Myopia, Hyperopia, Presbyopia



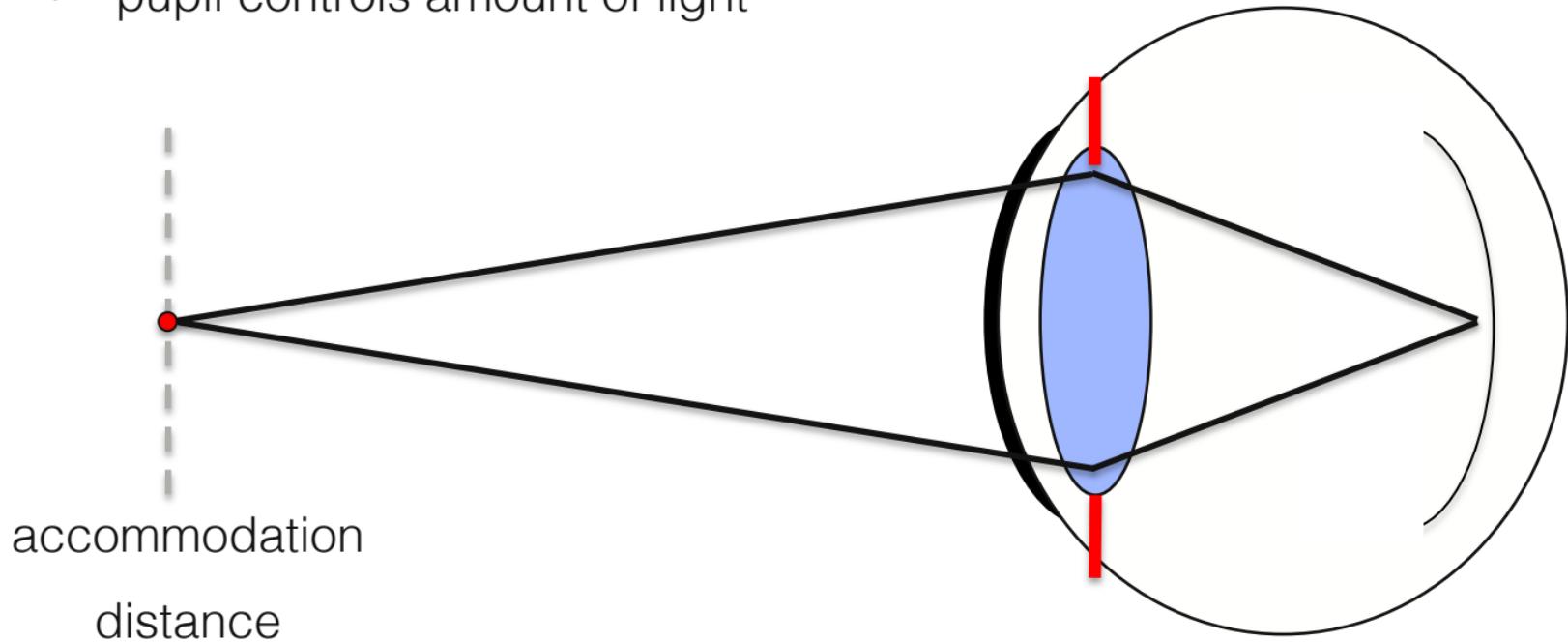
Retinal Blur

- pupil controls amount of light



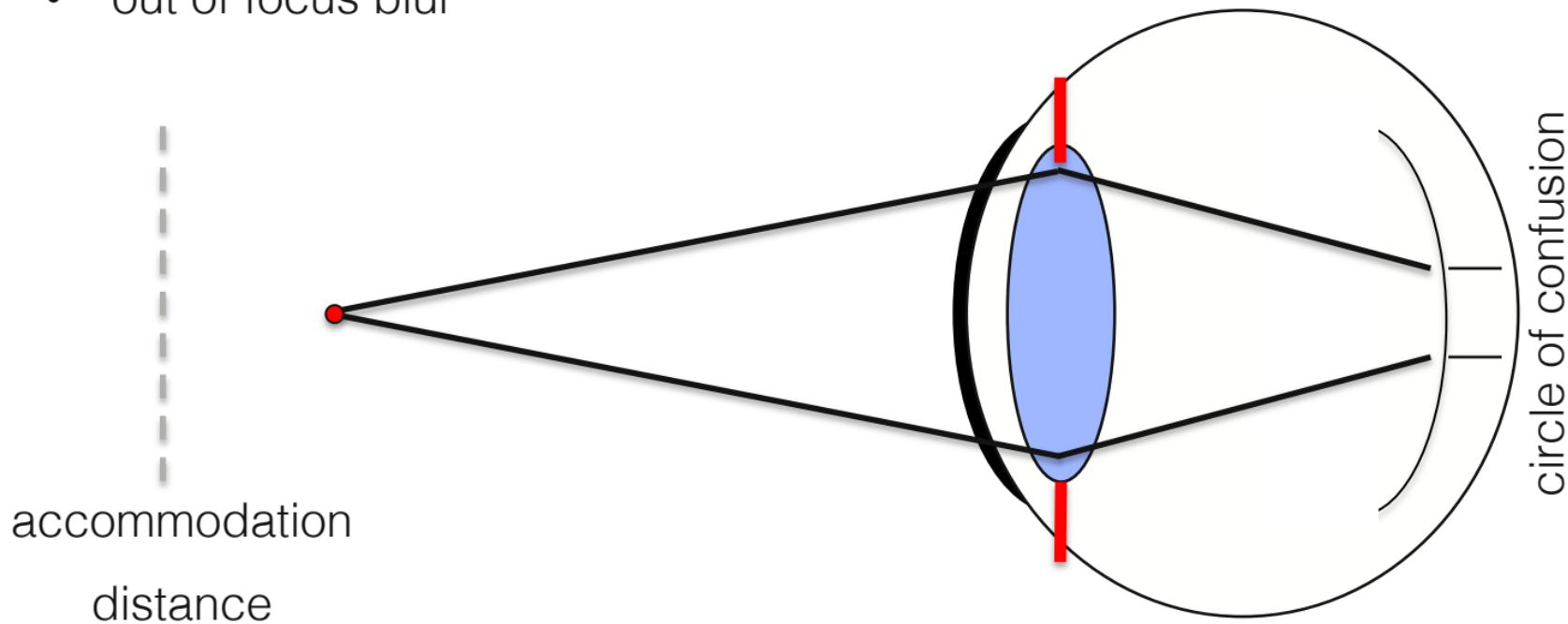
Retinal Blur

- pupil controls amount of light



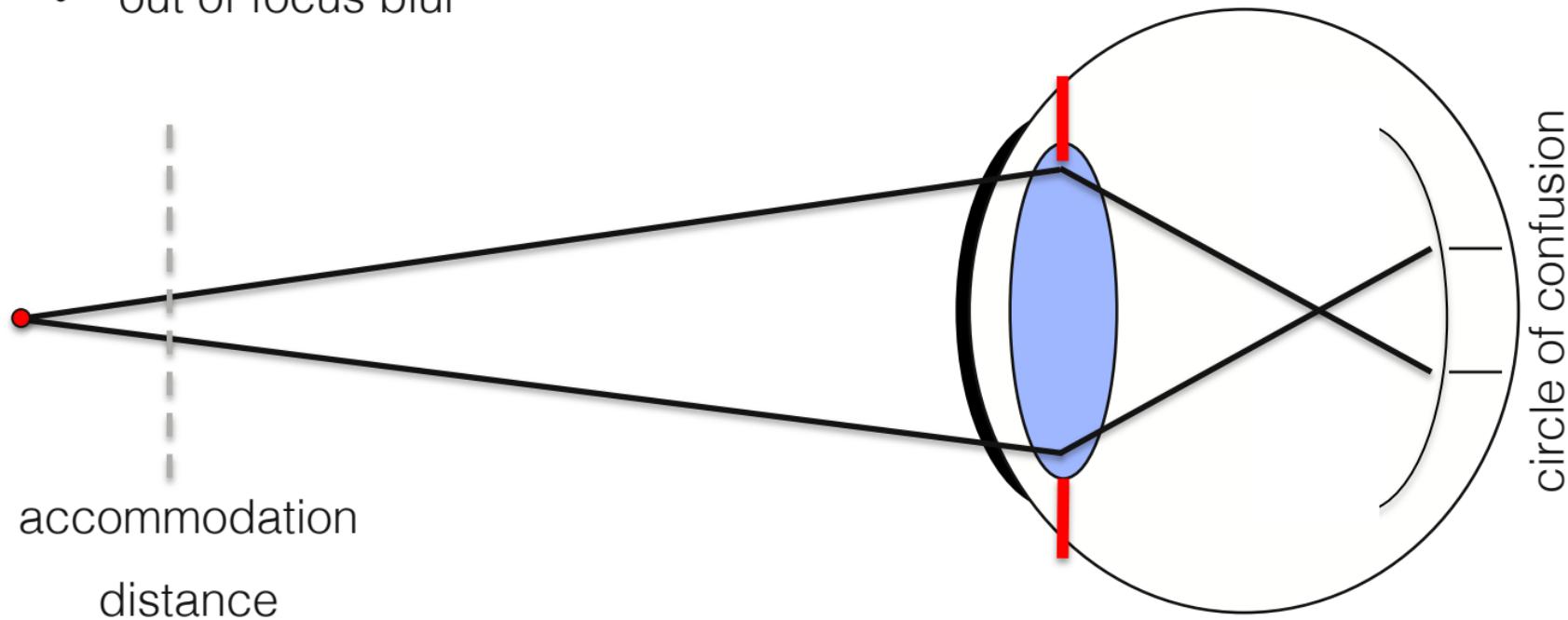
Retinal Blur

- out of focus blur



Retinal Blur

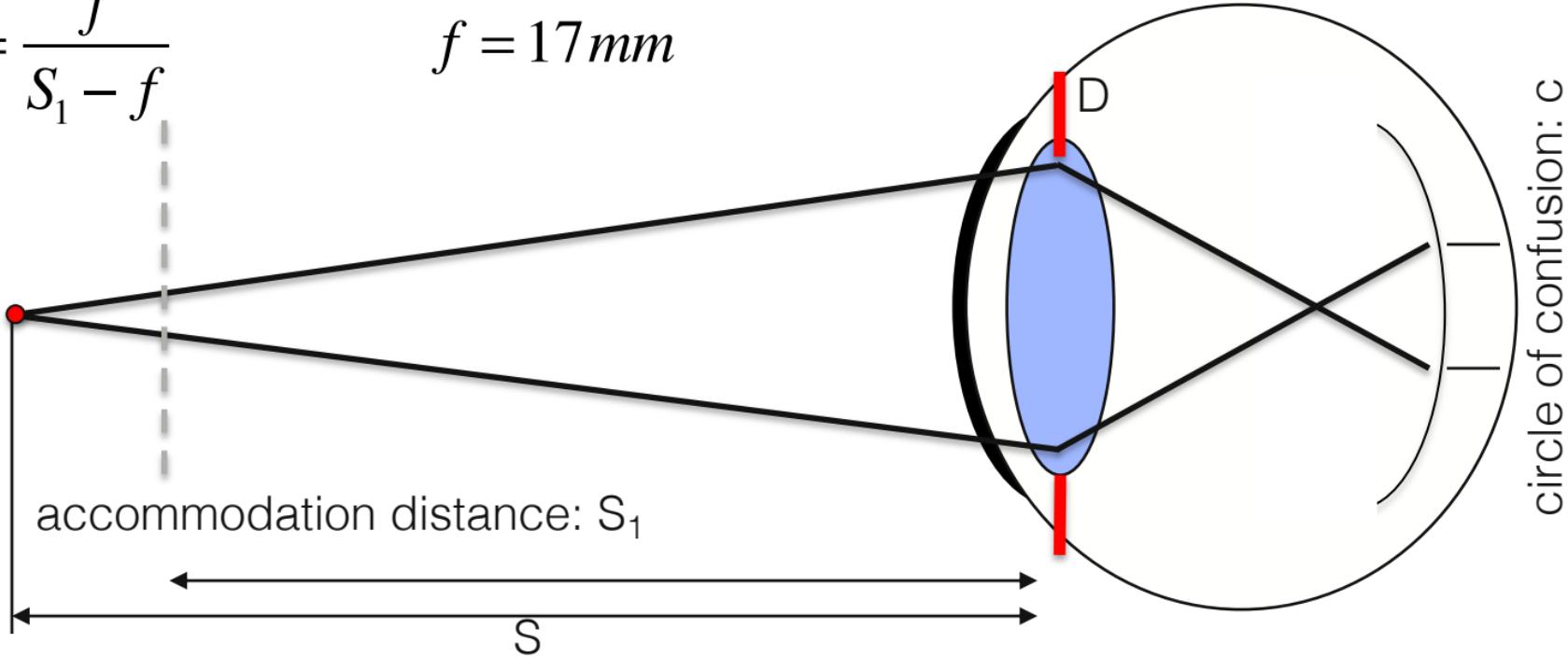
- out of focus blur



Retinal Blur / Depth of Field Rendering

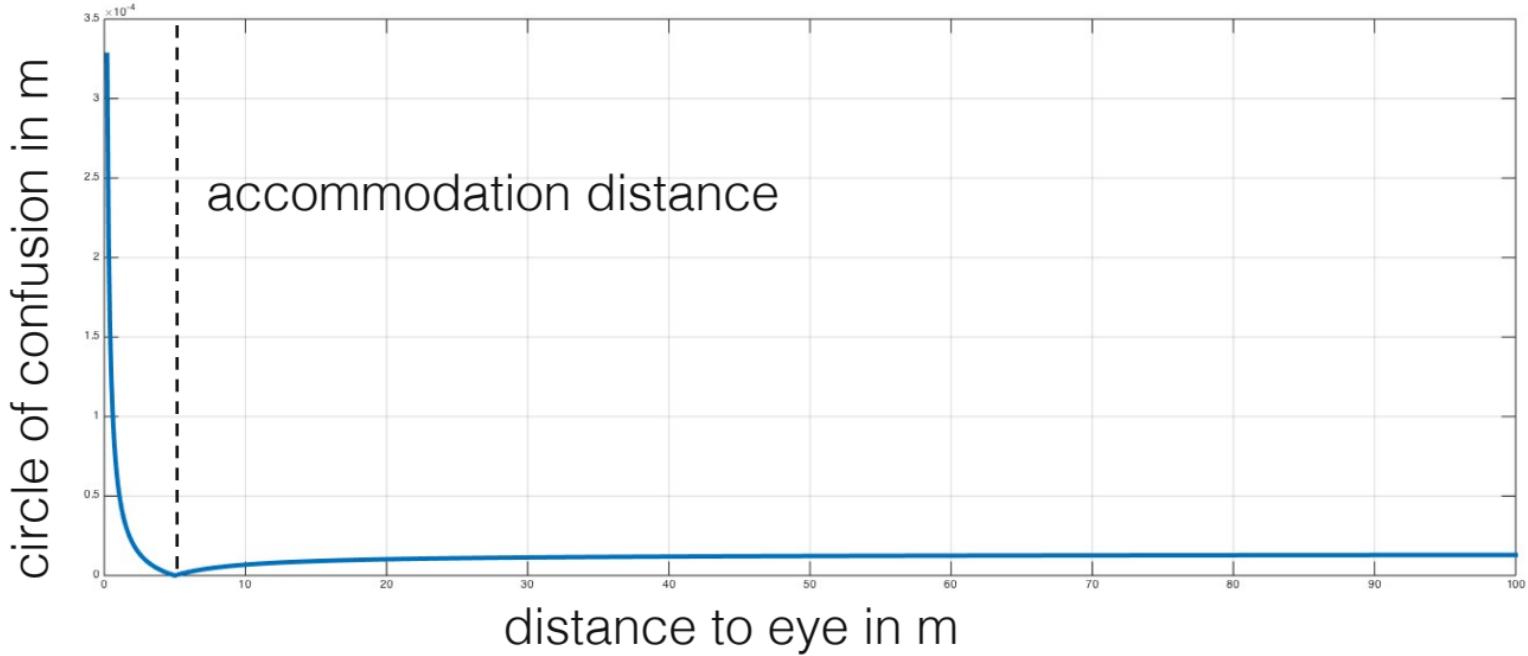
$$c = M \cdot D \cdot \frac{|S - S_1|}{S}$$

$$M = \frac{f}{S_1 - f} \quad f = 17\text{mm}$$



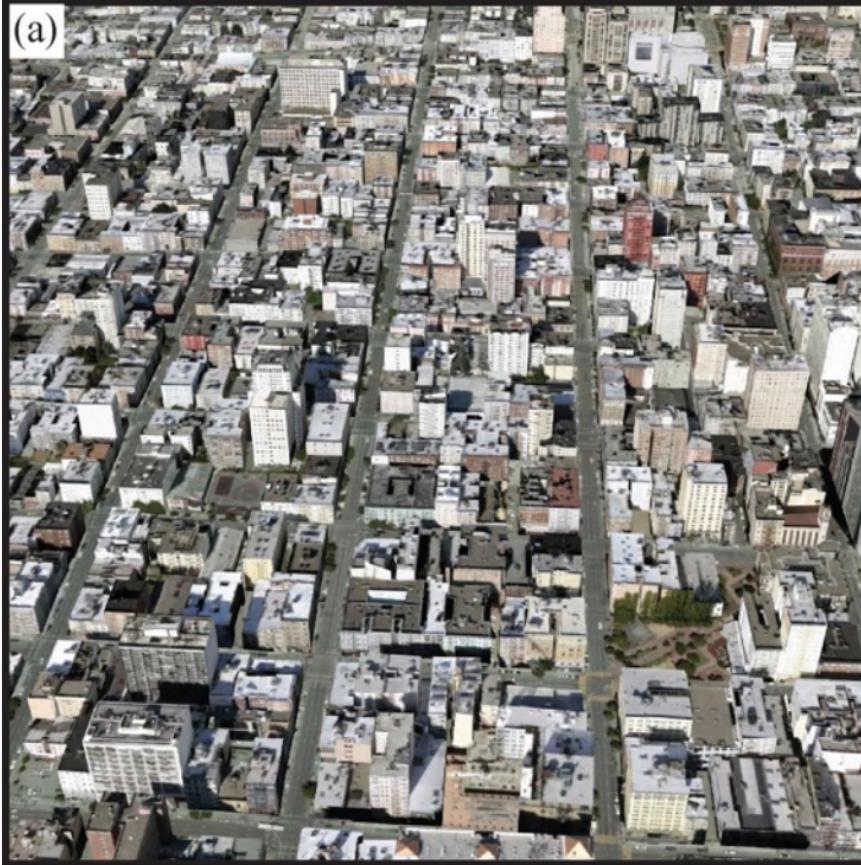
Circle of Confusion

$$c = M \cdot D \cdot \frac{|S - S_1|}{S}$$



Blur Affects Relative Object Size!

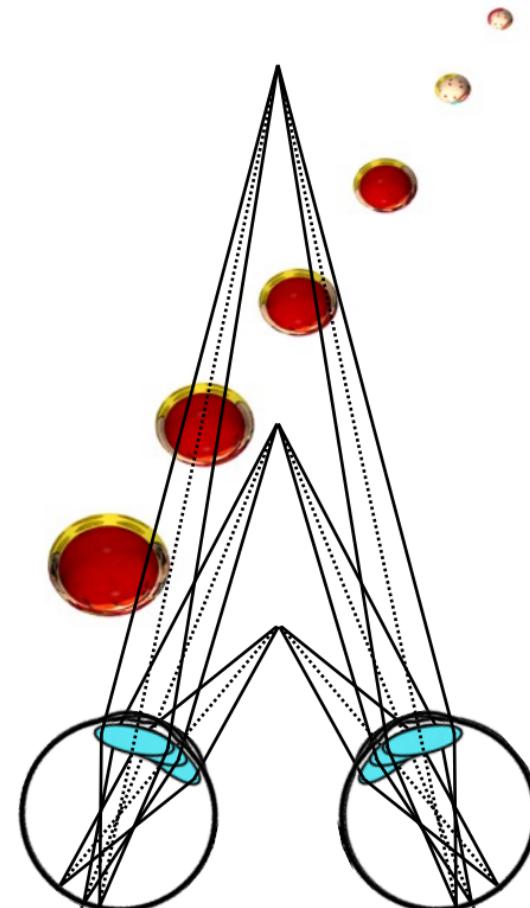
Held et al., 2006, ACM SIGGRAPH





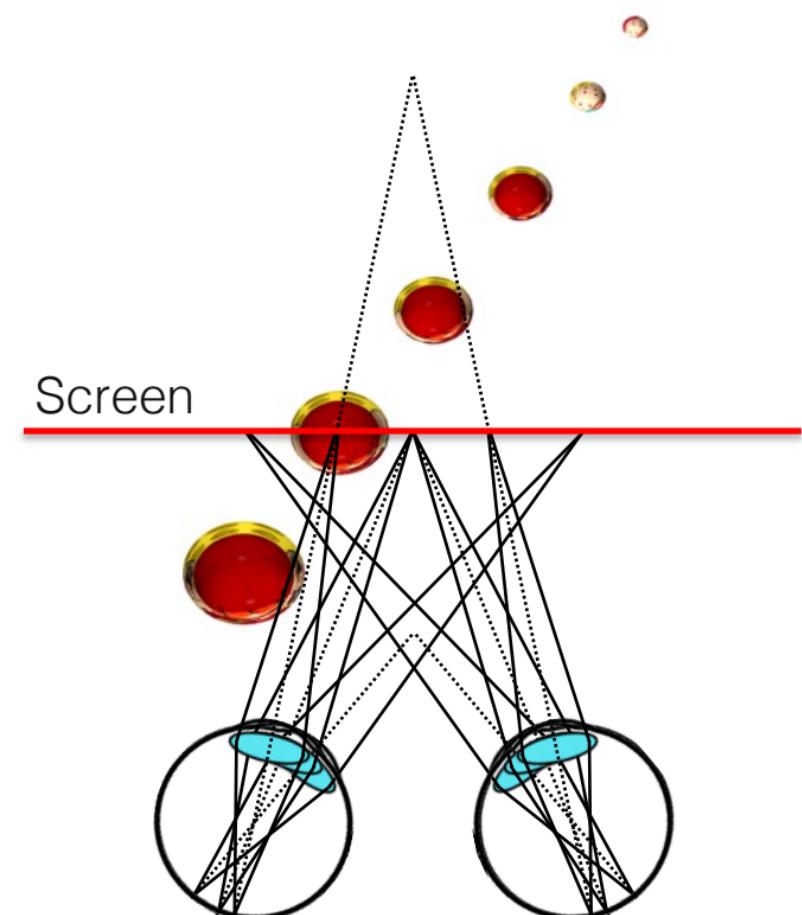
Top View

Real World:
Vergence & Accommodation **Match!**





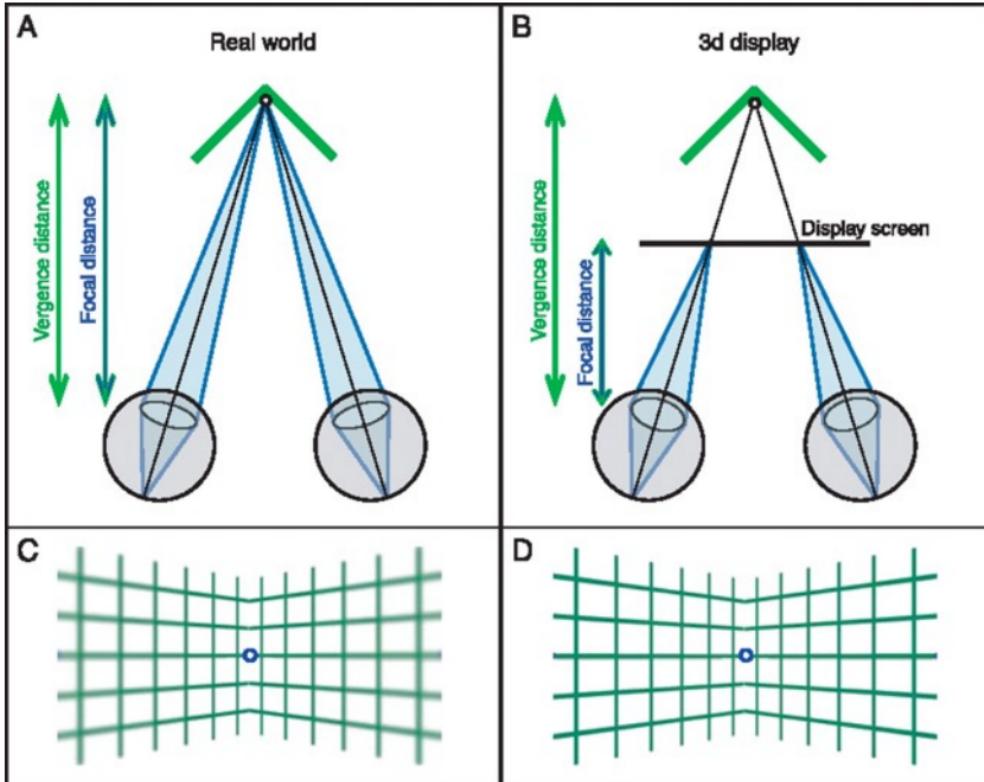
Top View



Stereo Displays Today:

Vergence-Accommodation **Mismatch!**

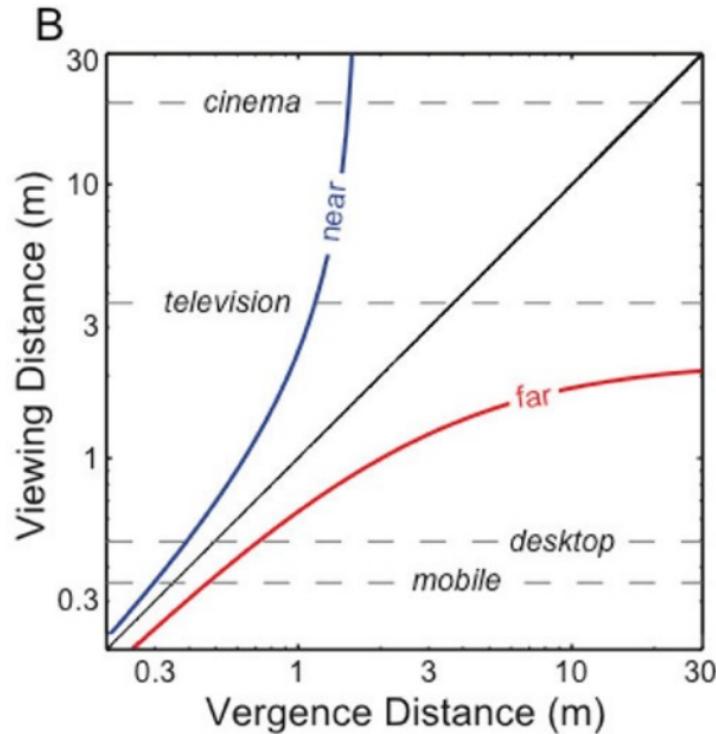
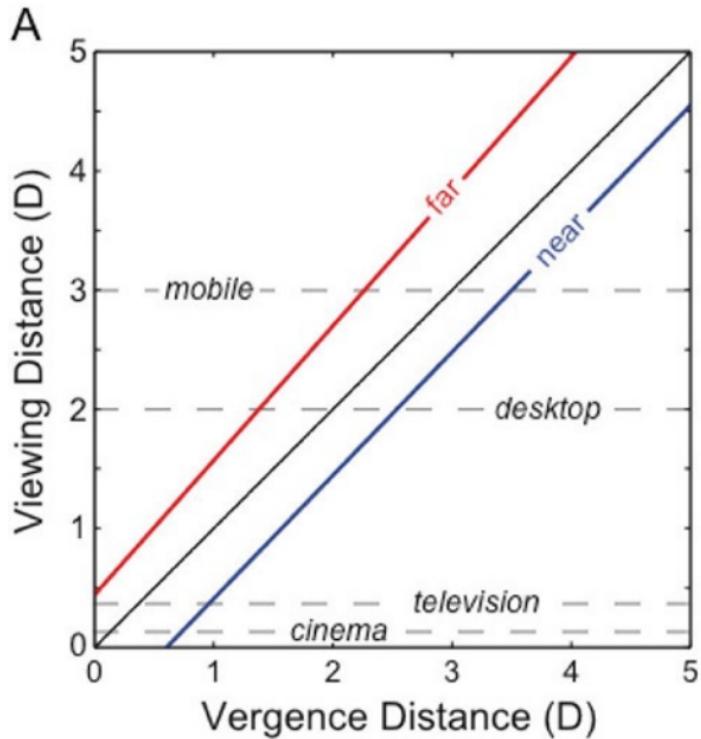
Vergence-Accommodation Conflict



effects

- visual discomfort
- visual fatigue
- nausea
- diplopic vision
- eyestrain
- compromised image quality
- pathologies in developing visual system
- ...

Zone of Comfort



Summary

- **visual acuity:** 20/20 is ~ 1 arc min
- visual acuity varies over retina: can exploit via **foveated rendering**
- **visual field:** $\sim 200^\circ$ monocular, $\sim 120^\circ$ binocular, $\sim 135^\circ$ vertical
- **temporal resolution:** ~ 60 Hz (depends on contrast, luminance)
- **depth cues in 3D displays:** disparity, vergence, accommodation, blur, ...
- **accommodation range:** $\sim 8\text{cm}$ to ∞ , degrades with age

References and Further Reading

interesting textbooks on perception:

- Wandell, "Foundations of Vision", Sinauer Associates, 1995
- Howard, "Perceiving in Depth", Oxford University Press, 2012

foveated rendering:

- Guenter, Finch, Drucker, Tan, Snyder "Foveated 3D Graphics", ACM SIGGRAPH Asia 2012
- Patney, Salvi, Kim, Kaplanyan, Wyman, Benty, Luebke, Lefohn "Towards Foveated Rendering for Gaze-Tracked Virtual Reality", ACM SIGGRAPH Asia 2016

depth cues and more:

- Cutting & Vishton, "Perceiving layout and knowing distances: The interaction, relative potency, and contextual use of different information about depth", Epstein and Rogers (Eds.), Perception of space and motion, 1995
- Held, Cooper, O'Brien, Banks, "Using Blur to Affect Perceived Distance and Size", ACM Transactions on Graphics, 2010
- Hoffman and Banks, "Focus information is used to interpret binocular images". Journal of Vision 10, 2010
- Hoffman, Girshick, Akeley, and Banks, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue". Journal of Vision 8, 2008
- Huang, Chen, Wetzstein, "The Light Field Stereoscope", ACM SIGGRAPH 2015

the retina, visual acuity, visual field

- Roorda, Williams, "The arrangement of the three cone classes in the living human eye", Nature, Vol 397, 1999
- Snellen chart: https://en.wikipedia.org/wiki/Snellen_chart
- Ruch and Fulton, Medical physiology and biophysics, 1960

contrast sensitivity function & hybrid images:

- Oliva, Torralba, Schyns, "Hybrid Images", ACM Transactions on Graphics (SIGGRAPH), 2006
- Spatio-temporal CSF: Kelly, Motion and Vision. II. Stabilized spatio-temporal threshold surface, Journal of the Optical Society of America, 1979